ESSAYS ON THE PRODUCTION OF HEALTH AND HEALTH CARE FINANCE

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of
Doctor of Economics

presented by
Johannes Schoder
from Esslingen a.N., Germany

approved at the request of
Prof. Dr. Peter Zweifel
Prof. Frank R. Lichtenberg, Ph.D.

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorises the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, 27.10.2010

the Dean: Prof. Dr. Dr. Josef Falkinger

Preface

In writing this dissertation, I have enjoyed the help and support of several people whom I would like to thank.

First and foremost, I am grateful to Prof. Dr. Peter Zweifel, my thesis supervisor, for his guidance and ongoing support. I appreciate all his contribution of time, ideas, and funding to make my doctoral thesis possible. In particular, I am grateful for the possibility to work in several related projects and the opportunity to study at the Columbia University. Furthermore, I want to thank Prof. Dr. Frank Lichtenberg for co-supervising this dissertation, his support during my studies at the Columbia University, and the exciting Squash matches at the Dodge Fitness Center. I am also indebted to Prof. Dr. Massimo Filippini for supervising one of my oral doctoral exams. Many thanks go to Prof. Dr. Bernd Raffelhüschen who introduced me to the field of health economics during my studies at the University of Freiburg (Germany).

A substantial part of the research leading to chapter 3 was conducted at the Columbia University. The financial support by the Swiss National Science Foundation (SNF) is gratefully acknowledged.

In addition, I would like to thank my former and current work colleagues at the Department of Economics (formerly the Socioeconomic Institute) of the University of Zurich and the Institute of Public Economics of the University of Freiburg (Germany). Dr. Stefan Fetzer, Dr. Jasmin Häcker, and Dr. Chrisitan Hagist (all University of Freiburg). Dr. Karolin Leuckert who strongly supported me to carry out my doctoral studies in Zurich, Angelika Brändle, Dr. Patrick Eugster, Dr. Boris Krey, and Dr. Ilja Neustadt for introducing me to the chair of Prof. Dr. Peter Zweifel. Thanks to Patrick's and Boris' cheerful manner it was easy to integrate, and especially due to Boris' encouraging and motivational skills I continued my doctoral studies. Furthermore I

want to thank my current collegues Philipp Morf, Maurus Rischatsch (for his ciritcal questions), Adrian Rohner, Micha Ruflin, PD Dr. Felix Schläpfer, Dr. Harry Telser, Maria Trottmann (for her excellent knowledge of Swiss health policy), and especially my office mate Philippe Widmer, who always supported me in questions related to latex, econometrics, and writing. I am in particular indepted to my co-author and colleague Michèle Sennhauser. Many thanks also go to Eva-Maria Aulich, for the inspiring disucssions during the lunch breaks, and especially Tanja Wasser. Last but not least, I want to thank the whole team of the library for their support.

My doctoral studies gave me the opportunity to meet wonderful people from all over the world, who also supported me in writing this thesis. During my time in New York special thanks go to Prof. Dr. Achim Schmitt and the many runs we had in central park. Marcos Agurto from the University of British Columbia and Prof. Dr. Steven Stillman, my skiing buddy from Wellington, New Zealand. Prof. Dr. Harald Barkhoff, University of Hawai'i at Hilo who showed me that everything is possible in life. Special thanks go to my friends Philip Burkert, Hannes Burkert, and Peter Weber.

Finally, this is also to my parents, Ellengard and Gustav Schoder and best brother I could imagine, Jörg Schoder, for all their love, encouragement, and continuous support during my doctoral studies. Their belief in my capabilities enabled me to complete this thesis.

Johannes Schoder, Zurich, February 2011

Contents

List of Figures

List of Tables

Chapter 1
Introduction

1 Introduction

This dissertation is a collection of four essays dealing with topics of the production of health and health care finance. In almost all industrialized countries public health care expenditure continues to increase at a faster pace than GDP. Thus, debate on health policy often focuses on limiting the growth of health care spending. However, from an economic perspective health care spending is justified as long as the marginal benefit exceeds or equals its marginal cost. Chapters 2 and 3 deal with topics of the production of health. While chapter 2 looks at the demand side of health, chapter 3 deals with the supply side of health. Chapters 4 and 5 turn to questions related to health care finance.

Taking a new perspective on the production of health, chapter 2 attempts to explain why countries are still spending on health although there is evidence of decreasing marginal returns w.r.t health care expenditure (so called "flat-of-the-curve medicine"; see e.g. Enthoven, 1980). Tracing a panel of 24 OECD countries over time, chapter 2 finds that uncertainty with regard to time to death (as measured by the standard deviation of age at death) has declined over the past four decades. Assuming that individuals are risk-averse with regard to their health, this reduction of uncertainty may well justify health care spending. Estimating a modified production function, chapter 2 finds that health care expenditure has significantly reduced the standard deviation of age at death. Furthermore, willingness to pay for such a risk reduction exceeds marginal health care expenditure of the United States and Switzerland, implying that additional health care expenditure may be worth its cost.

Chapter 3 determines whether there are economies of scale in the production of health. While health was measured on the aggregated country level in chapter 2, micro-level data is used in chapter 3. Based on the observation that 5-year survival rates for breast cancer vary by more than 10 percent across U.S. regions, the factors contributing to

the different survival rates are analyzed. The results indicate that patients tend to survive longer (than the US average) in those regions where relatively more cancers of the same type exist. One likely explanation for this finding is, that the relatively higher prevalence of cancers in some regions has lead to greater accumulation of disease-specific knowledge and experience which translated into improved health outcomes.

However governments facing the pressure of public health care expenditure not only just limit the growth of health care expenditure but also seek ways to supply health care efficiently. Chapter 4 deals with health care reforms that attempt to increase competition between providers of health care and health insurers in order to raise the overall performance of health care systems. One finding of this chapter is that increasing competition also requires an appropriate institutional setting such as an adequate risk adjustment. However, as chapter 5 shows, refining the risk adjustment scheme turns out to be a difficult task triggering unintended side effects which in turn may reduce the overall performance of a health care system.

In the Netherlands, Germany, and Switzerland Managed Care has been one of the preferred options to increase the performance of their health care system. By vertically integrating health insurance and health care provision, Manged Care attempts to improve the allocation of resources in health care while limiting health care expenditure. Chapter 4 evaluates elements of Managed Care of these three insurance-based countries with regard to their contribution to the performance of the entire health care system. The country comparison reveals that Managed Care depends stongly on the institutional setting. The more freedom to contract between consumers, health insurers, and health care service providers, the greater its contribution to the health care system.

Even though most of the literature emphasizes the effect of Managed Care as a device to improve the allocation of resources in health care, health insurers can also use Managed Care to attract favorable risks or to deter unfavorable ones. As a consequence, Switzerland is refining its risk adjustment formula, including a third indicator (beside age and gender) “hospitalization of more than three days in the previous year”. Chapter 5 addresses the consequences of this refinement for an individual Swiss health insurer. Due to its high share of Managed Care contracts (and therefore low hospitalization rates of its insureds), payments into the risk adjustment

scheme are predicted to explode. The likely response of the insurer's risk management is to extend hospital stays beyond three days, which runs counter stated policy objectives. While the fine-tuning of the risk adjustment scheme decreases risk selection efforts, it likely discourages health insurers applying Managed Care in general. Thus, cost savings through Managed Care that are not attributable to risk selection efforts will be lost as well, resulting in a further increase of public health care expenditure.

Note that Peter Zweifel co-authored chapters 2, 4, and 5, Frank Lichtenberg co-authored Chapter 3, and Michèle Sennhauser co-authored Chapter 5. Chapter 4 appeared in the *Swiss Journal of Economics and Statistics*. Chapter 5 is published in the *International Journal of the Economics of Business*.

Johannes Schoder
Zurich, February 2011

Chapter 1

Flat-of-the-curve Medicine: A New Perspective on the Production of Health

Johannes Schoder and Peter Zweifel

Submitted to

Health Economics Review

2 Flat-of-the-curve Medicine: A New Perspective on the Production of Health

2.1 Introduction and motivation

Industrial countries have been spending a rising share of their economic resources on health care. From 1960 to 2004 health care expenditure (HCE) of OECD countries increased from 3.8 percent of GDP to 8.9 percent on average. Over the same period, health outcomes measured by average life expectancy at birth improved from 68.4 to 78.5 years. However, this increase has slowed recently. In the United States e.g., it has been 0.19 percent p.a. between 1980 and 2004, down from 0.3 between 1960 and 1980. Since HCE continued to grow at a rate of 7.7 percent p.a. between 1980 and 2004, this has often been interpreted as evidence of decreasing marginal returns ("flat-of-the-curve medicine"; Enthoven, 1980, Fuchs, 2004), raising the question of why citizens and governments failed to reallocate resources away from medicine.

However, calling for such reallocation may be premature on at least two accounts. First, several studies find that marginal returns to HCE still outweigh its marginal cost (Cutler and McClellan, 2001, Murphy and Topel, 2006, Lichtenberg, 2007a). This would explain why, in countries where individual willingness to pay for medical services tends to prevail over political considerations of cost control (such as the United States, the Netherlands, and Switzerland), the share of HCE in the GDP keeps increasing. Second, the implicit assumption that individuals only value changes in the expected value of health status is open to criticism. If people are risk-averse with regard to their health, they value a reduction in the variance of health status even if its expected value does not change. Thus, judging the benefits of HCE by its marginal product in terms of expected health (as traditionally done in studies of the production of health) possibly neglects the willingness to pay of risk-averse individuals

for reduced uncertainty surrounding their health status.

Following up on this second aspect, this study seeks to determine whether the marginal cost in terms of HCE matches its marginal benefit if the reduction in uncertainty surrounding life expectancy is accounted for. In order to do this, we will proceed as follows. First, a conventional production function with life expectancy as the dependent variable is reestimated to verify that the countries of our sample are characterized by flat-of-the-curve medicine. Second, we examine whether a reduction of uncertainty surrounding life expectancy indeed occurred over the past 46 years. Third, based on the econometric estimation of an appropriately modified health production function we determine the relative contribution of medical and non-medical inputs to reduced uncertainty. Finally, we compare the marginal cost of HCE with its marginal benefit in terms of willingness to pay for reduced uncertainty.

We find HCE as well as GDP to be significant determinants of the variance of health status. A 10 percent increase of HCE is estimated to lead to a 0.42 percent reduction of the standard deviation of life years. Furthermore, according to our calculations willingness to pay both in the United States and Switzerland for such a reduction exceed the extra HCE, implying that additional HCE may be worth its cost as "real insurance", bringing back health status to normal when illness strikes. Hence, flat-of-the curve medicine need not be wasteful.

Our study is closely related to the empirical literature on the production of health (e.g. Auster et al., 1969, Miller and Frech, 2000, Thornton, 2002, Shaw et al., 2005, and Zweifel et al., 2005). However, rather than assessing the contribution of inputs exclusively to the expected value of health status, it estimates their impact on the variability of health status as well. Furthermore, this work complements studies conducted on the individual level deriving willingness-to-pay values for health risk reductions either from experiments (see e.g. Cameron et al., 2010) or from utility-theoretic models (see e.g. Edwards, 2008).

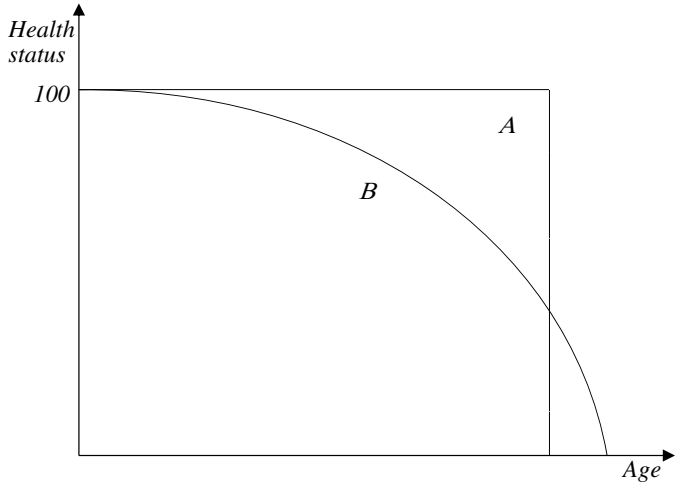
The remainder of this paper is structured as follows. The general background is presented in Section 2.2. Section 2.3 is devoted to the description of the data. Section 2.4 contains the econometric specification. Estimation results are presented and discussed

in Section 2.5. Section 2.6 provides estimates of the willingness to pay for more certain health. Section 2.7 concludes with a summary of key findings.

2.2 Risky health: Valuation and measurement

The basic hypothesis underlying this work is that individuals have preferences with regard to health profiles that are reflected in survival curves and their development over time. As a convenient starting point, consider the two hypothetical health profiles of Figure 2.1. Health profile A presumably represents the ideal of western lifestyle, living in perfect health followed by sudden death, indicated by a health status of zero (Fries, 1980). In contrast, health profile B represents an alternative where health status deteriorates with age but remains positive up to a higher age, indicating survival. The two profiles can also be interpreted as reflecting the probability of being in perfect health, which starts at 100 percent and stays there (profile A) or decreases with age (profile B). Thus, they represent cumulative distribution functions (cdfs), defined over the absence of death rather than (say) wealth. Even if profile B should have higher expected value, they can still be ranked in terms of second-order stochastic dominance (see e.g Laffont (1999), ch. 2.5). In the present case, the triangle-like area above profile B (indicating the cumulative difference between the two cdfs) exceeds the extra area below profile B. In this event, an individual who is risk-averse with regard to health status prefers profile A. He or she has a positive willingness-to-pay (WTP) for living in a country with a health profile A rather than B.

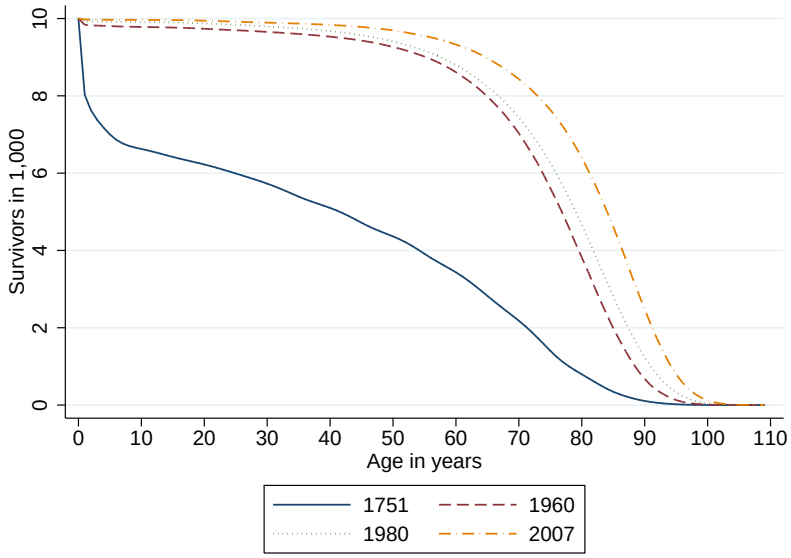
Figure 2.1: Ranking of two health profiles



Health profiles of this type are not available at this time. However, if individuals are successful in moving from profile B to the more rectangular profile A, they should in the aggregate exhibit an increasingly rectangular survival curve because survival constitutes the necessary (but not sufficient) condition for being in perfect health. Therefore, variability of age at death will be used as an indicator of uncertainty surrounding health status.

Various indicators of variability of age at death (VAD) are used in the literature such as the interquartile range (Wilmoth and Horiuchi, 1999), the Gini coefficient (Shkolnikov et al., 2003 and Peltzman, 2009), and the standard deviation (Kannisto, 2000). Regardless of choice of indicator, these studies document a secular decline in VAD for industrial countries, albeit at a somewhat reduced rate during the most recent decades. This development is tantamount to a rectangularization of the survival curve (see Figure 2.2 showing the case of Sweden). In keeping with the argument above, it is interpreted as evidence of individual's improved control over their health status. Note also that this improvement increasingly is reflected in the neighborhood of the nearly vertical segment of the nearly rectangular survival curve, calling for a special focus on the VAD of the elderly.

Figure 2.2: Rectangularization of the survival curve in Sweden



Source: Human Mortality Database (2008).

What this evidence is silent about is how individuals might have achieved this added control. In analogy with the production of health literature, it would be interesting to know whether the major contribution came from medical or non-medical inputs. To the authors' knowledge, there is only one study that relates a measure of VAD to medical and non-medical inputs (Le Grand, 1987). Due to data limitations, Le Grand (1987) performs but a cross-sectional regression for 1982 including 17 OECD countries. He relates public HCE, total HCE, GDP per capita, and a measure of income inequality (the share of the bottom quintile in national income) to the Gini coefficient of mortality. Public HCE has the expected negative impact on VAD but remains insignificant. Surprisingly, total HCE has a positive impact, whereas a higher GDP per capita and less income inequality are associated with lower VAD. However, the Gini coefficient is not translation independent (i.e. it changes if mean age at death differs between two periods or countries although the absolute differences between individuals' ages at death are the same). As to the interquartile range, it violates the transfer principle (it ignores a change in the distribution of deaths within a given age class if the number of deaths in that class does not change). For this reason, the standard deviation will be used as an indicator of VAD in the analysis below.

The present study extends previous work in three ways. First, it uses panel data tracing 24 OECD countries[1] over the past 45 years, permitting to test the robustness of the results found by Le Grand (1987). Second, since the relative contribution of medical and non-medical inputs may well change with age, VAD among the elderly (where rectangularization of the survival curve has been especially marked, see Figure 2.2) is examined in particular. Third, using evidence on the willingness to pay (WTP) for health risk reduction, WTP values are calculated for the two countries with the highest HCE per capita, the United States and Switzerland. These values are compared with the extra HCE to determine whether flat-of-the-curve medicine may be still worthwhile thanks to its effect on VAD. However, to address these research questions we first have to determine whether the countries of our sample indeed operate (on average) on the flat-of-the-curve. In sum, this leads to the following four research questions:

Q1: Are the countries of our sample characterized by flat-of-the-curve medicine?

[1]These are Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, Germany, Hungary, Iceland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Slovak Republic, Spain, Sweden, Switzerland, United Kingdom, and the United States.

Q2: Do medical or non-medical inputs contribute more to reducing variability of age at death (VAD)?

Q3: Are these effects different for VAD among the elderly?

Q4: Is flat-of-the-curve medicine wasteful?

2.3 Data

Data to compute the dependent variables used below (remaining female life expectancy at age 60 and two measures of variability of age at death) are obtained from the Human Mortality Database (2008) (HMD). The Human Mortality database provides two different variants of life tables, cohort and period life tables. The former represent the mortality experience of individuals who are born in the same year and thus are truly comparable. However, cohort life tables contain complete mortality information only on cohorts without any survivors left. By way of contrast, period life tables show the estimated number of survivors at age x if a hypothetical birth cohort of 100,000 born today have the mortality rates that are observed today for people at various ages up to x (Human Mortality Database, 2008).[2]

Therefore remaining female life expectancy at age 60 (LEF_{60}) and the standard deviation of age at death (sd) are calculated from life tables as follows. LEF_{60} is the weighted average of age at death above 60, $x - 60$, with weights given by the number of females dying at the respective age, d_x, relative to the number of survivors at age 60, l_{60}

$$LEF_{60} = \sum_{x=60}^{\omega} \frac{d_x \cdot (x - 60)}{l_60} = \sum_{x=60}^{\omega} f_x \cdot (x - 60), \tag{2.1}$$

where ω the maximum age in the life table. This variable will be used to check whether industrial countries indeed are on the flat-of-the-curve with regard to HCE (see Section 2.5.1 below).

The overall sd of age at death is given by

$$SD = \sqrt{Var} = \sqrt{\sum_{x=0}^{\omega} f_x (x - le)^2}, \tag{2.2}$$

[2] A disadvantage of period life tables is that they are based on one-year age intervals.

with *le* symbolizing life expectancy at birth. In keeping with Figure 2.1, perfect rectangularization means a vertical drop of the health (and hence survival) profile. The age at which the greatest number of a cohort's members die approximates best this vertical drop. Therefore, the standard deviation above the mode will be used to measure compression of mortality, which increasingly occurs at higher ages (see the example of Sweden again in Figure 2.2). The sd above the mode is calculated in analogy to the overall sd,

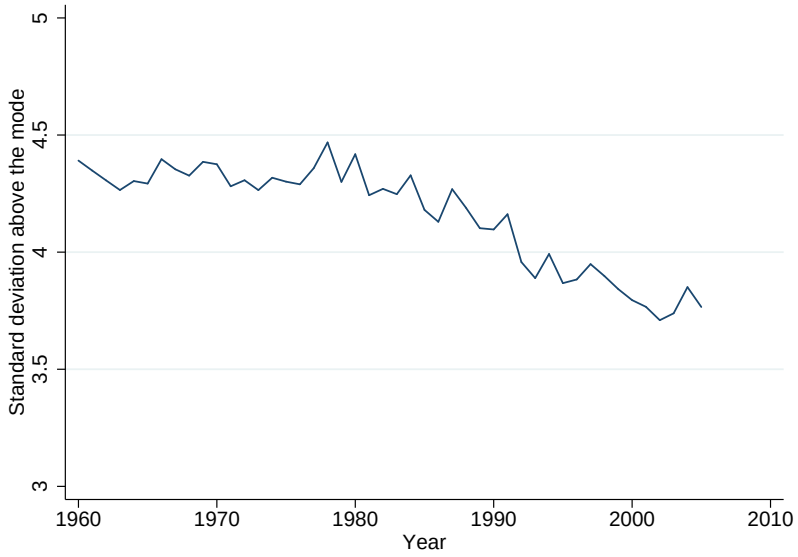
$$SD_{mode} = \sqrt{\sum_{x=mode}^{\omega} f_x(x - mode)^2}. \tag{2.3}$$

To test the robustness of the sd measure, we provide results for the Gini coefficient in the Appendix. The overall sd and the Gini coefficient exhibit a similar downward trend between 1960 and 1995 when averaged over 24 OECD countries[3] see Figures 2.5 and 2.6 of the Appendix. Figures 2.7 and 2.8 of the Appendix trace the overall sd and the Gini coefficient for six selected countries. The pattern of decline confirms the similarity between the two measures. Interestingly, the ranking between countries changed over time. In 1960, the Italians, Portuguese, and Japanese faced higher uncertainty of age at death than U.S. citizens (up to 7 years in terms of sd). However, by 2005 Americans faced considerably higher VAD than the citizens of these countries (almost 3 years in terms of sd).

The sd above the mode exhibits a less regular pattern than both the overall and the Gini coefficient, mainly due to shifts in the age with maximum mortality. It remains roughly constant in the 1960s and 70s but has declined since 1980 (see Figure 2.3 below). Figure 2.9 of the Appendix shows that rectangularization at higher ages differs considerably between the six selected countries. It has declined for all six countries (except for the United States); however, the pattern of decline is again less regular than that of the overall sd and the Gini coefficient.In sum, the indicators for VAD confirm the findings of previous studies (see Section 2.2). They suggest that individuals in industrialized countries have been exposed to less uncertainty regarding their longevity (and presumably health status) since the 1960s, although differences between countries and subperiods (especially for the sd above the mode) persist. The more the question of what may have contributed to these differential developments gains importance.

[3]The correlation between the two measures is 0.97.

Figure 2.3: Standard deviation above the mode averaged over 24 OECD countries, 1960 to 2005



Source: Human Mortality Database (2008).

We defined medical and non-medical inputs drawing on the literature of the production of health (Auster et al., 1969, Miller and Frech, 2000, Thornton, 2002, Shaw et al., 2005, and Zweifel et al., 2005). Due to missing values in the OECD health data base only the following determinants are retained. These are GDP per capita and alcohol consumption per capita in liters for the non-medical and health care expenditure per capita for the medical input. The OECD data is known for some problems. One of them is national differences with regard to the delimitation of the health care sector, resulting in different baskets of services, another the lack of comparability and precision of health care deflators. In the case of countries such as Switzerland or the Netherlands, HCE covered by basic health insurance are termed private HCE although basic insurance is mandatory and regulated by the government.[4] In view of these difficulties, HCE is not split into private and public HCE (contrary to Le Grand, 1987). Furthermore, HCE is not deflated using national price indexes but by the exchange rate when converting the figures into USD, thus avoiding PPP indicators that may contain additional meassurement error (see Gerdtham and Jönsson, 1991).

[4]E.g. premiums for the basic mandatory coverage are not risk-rated.

The sample comprising the 24 countries is characterized in Table 2.1 below. As to sd as the indicator of VAD, it decreases from 19.34 years in 1960 to 14.74 years in 2005. Decomposition of sd suggests that within-country differences (sd_w) are rather more important than between-country ones (sd_b), indicating that decreases over time are the primary source of variation. Turning to the independent variables, one observes slower growth of HCE and GDP in recent years. Whereas total HCE and GDP per capita increased by factors of 13 and 8, repectively between 1960 and 1983, these factors decreased to 4.5 and 3.5 between 1983 and 2005.Interestingly, alcohol consumption per

Table 2.1: Descriptive statistics of variables, selected years

Variable	Mean	1960	1983	2005	sd_o	sd_b	sd_w	N
LEF_{60}	21.64	18.84	21.38	24.54	2.11	1.22	1.75	997
SD	16.81	19.34	16.44	14.74	1.76	0.97	1.49	1,102
GINI	11.55	13.73	11.24	9.68	1.63	0.92	1.35	1,101
SD_{mode}	4.17	4.34	4.33	3.81	0.39	0.19	0.35	1,102
HCE	1,283	60.25	792.28	3,436	1,187	636.22	1068	838
GDP	13,866	1,341	10,287	35,782	11,903	7,787	10,780	965
ALC	10.62	7.87	11.81	9.41	3.66	3.27	1.78	1,003

Note: Gini is multiplied by 100. The countries included are Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, Germany, Hungary, Iceland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Slovak Republic, Spain, Sweden, Switzerland, United Kingdom, United States.

capita (ALC) increased first but has been diminishing after reaching a peak in the mid 1970s. As to the decomposition of the standard deviation of the independent variables, variation over time again exceeds the between standard deviation (sd_b), except for alcohol consumption.

2.4 Econometric specification

One way to assess the contribution of a set of inputs to remaining female life expectancy at age 60 (LEF_{60}) and variability of age at death (VAD) is by eliminating certain causes of death from the data (see Shkolnikov et al., 2003, Wilmoth and Horiuchi, 1999, but also Andreev, 1982, and Lamber and Aronson, 1993). However, comparing different countries over time entails the problem that this contribution may be conditioned by country-specific characteristics (e.g. the type of health care system). In contrast, econometric techniques designed for panel data permit to control

for heterogeneity between countries either through fixed or random effects. In the fixed effects (FE) approach, the country-specific effects, c_i, are included in the set of independent variables as a set of country-specific dummies. Alternatively, the c_i can be netted out by measuring all variables as differences from their country-specific means. The random effects (RE) approach assumes the c_i to be stochastic, which means they must be uncorrelated with the independent variables for unbiased parameter estimation. Both the RE and FE estimation were found to suffer from heteroskedasticity, reflecting cross-sectional correlation of error terms in Eqs. (2.4) and (2.5) below. Correcting for heteroskedasticity with a first-order autoregressive error term [AR(1) process] and applying the Hausman test we find that RE is preferred over FE throughout at the 5 percent significance level or better.

However, two additional issues need to be clarified. First, medical and non-medical inputs were found to influence remaining life expectancy with a lag by Zweifel et al. (2005). The same may be true for our sample. Alcohol consumption, for instance, likely does not undermine control over one's health immediately but rather over the course of years. Likewise, earlier HCE may also contribute to higher LEF_{60} and lower VAD, repsectively. As to GDP, it is interpreted as representing the budget contraint rather than an indicator of life style, which argues against the introduction of a lag. Based on the Hausman test, we choose an optimal lag length of 10 years for HCE and ALC in Eq. (2.4) and 5 years for HCE in Eq. (2.5), values that seem to be reasonable in view of earlier research. The other issue is endogeneity. LEF_{60} and VAD may feed back to HCE. Countries where individuals live shorter or face higher uncertainty with regard to longevity may spend more on HCE than countries where individuals live longer or face less uncertainty. Such a feedback would likely occur through the political process, in analogy to the feedback relationship found by Zweifel et al. (2005). However, the Durbin-Wu-Hausman test (Durbin, 1954 and Hausman, 1978) for endogeneity does not reject the null hypothesis of exogeneity of HCE at the one percent level.

Based on the econometric specification of Le Grand (1987) and the conventional health production approach (see e.g. Zweifel et al., 2005) the following specification is estimated [note that the variables in Eq. (2.4) are in logarithms]:

$$\begin{aligned} LEF_{60it} &= \alpha_0 + \alpha_1 HCE_{it-10} + \alpha_2 HCE_{it-10}^2 + \alpha_3 GDP_{it} + \alpha_4 GDP_{it}^2 \\ &\quad + \alpha_5 ALC_{it-10} + \alpha_6 ALC_{it-10}^2 + c_i + \gamma_t + u_{it}. \end{aligned} \tag{2.4}$$

$$\begin{aligned} VAD_{it} &= \beta_0 + \beta_1 HCE_{it-5} + \beta_2 HCE_{it-5}^2 + \beta_3 GDP_{it} + \beta_4 GDP_{it}^2 \\ &\quad + \beta_5 ALC_{it-10} + \beta_6 ALC_{it-10}^2 + c_i + \gamma_t + u_{it}. \end{aligned} \tag{2.5}$$

The dependent variables are,

- LEF_{60}: Remaining female life expectancy at age 60.
- VAD: Variability of age at death of country i in year t measured by the overall sd and the sd above the mode, calculated according to Eqs. (2.1), (2.3).

The independent variables are,

- HCE_{-5}: Total private and public HCE per capita, nominal but converted in 1,000 USD. Devoting more resources to health care is expected to enhance control over health status, and hence higher LEF_{60} and lower VAD, repectively. Therefore, α_1 is predicted to be poitive and β_1 to be negative.
- GDP: GDP per capita in 1,000 USD, nominal but converted in 1,000 USD. This variable first of all reflects the budget contraint. Now, length of life and control over one's health status is quite likely a normal good, the demand for which increases with average income, ceteris paribus. Second, however, average income is importantly determined by labor productivity. To the extent that non-market and market productivity develop in a similar way, a higher value of GDP reflects a population that is better able to control their health status. In this way, GDP also serves as an overall indicator of non-medical inputs. Both arguments suggest a positive sign for α_3 and a negative sign for β_3.
- ALC_{-10}: Annual consumption of pure alcohol in liters per person above the age of 15. Lower values indicate a healthier lifestyle implying improved health and control over health status. Hence, α_5 is predicted to be positive and β_5 to be negative.
- c_i: A set of country-specific dummies.
- γ_t: A set of year-specific dummies to control for a possible time trend.
- u_{it}: A stochastic error term, assumed to be i.i.d. normal.

2.5 Estimation results

2.5.1 Checking for flat-of-the-curve medicine

Table 2.2 presents RE estimation results for (arithmetic) LEF_{60}. This choice of dependent variable permits a comparison with Zweifel et al. (2005), who estimated the same specification, but using a different sample[5]. In general, the estimated coefficients roughly correspond with these earlier estimates values (see column entitled *ZSE* of Table 2.2). In keeping with these earlier estimates, HCE exhibits decreasing marginal returns.

Table 2.2: Determinants of remaining life expectancy for females at age 60, 1960-2005

LEF_{60}	Coef.	z	P>z	Coef. $(ZSE)^a$
HCE_{-10}	1.143	3.4	0.001	2.045**
HCE^2_{-10}	-0.341	-2.78	0.005	-0.565**
GDP	0.0602	3.01	0.003	0.122**
GDP^2	-0.001	-2.99	0.003	-0.004**
ALC_{-10}	-0.060	-2.63	0.298	-0.043
ALC^2_{-10}	-0.002	-1.16	0.245	0.002
constant	19.537	0.394	49.64	18.57
rho_{ar}	0.904			
Wald χ^2	714.24			
Prob>χ^2	0.904			
R-squared	0.4976			
Observations	631			

Note: **p<0.01.[a]Estimates from Zweifel et al. (2005).

With regard to remaining life expectancy, the critical value of HCE beyond which its marginal effect ceases to be positive can be put at USD 1,675.[6] With a mean value of USD 3,436 as of 2005 OECD countries on average are well within the flat-of-the-curve range. Therefore, as to research question Q1 stated in Section 2.2, we can conclude

[5]They included consumption of kilocalories per capita as an additional lifestyle variable, which however turned out to be not significant.

[6]From Table 2.2, one obtains the critical value beyond which *e(LEF, HCE)* decreases: $\frac{\partial LEF}{\partial HCE} = 1.143 - 2 \cdot 0.341 HCE = 0$. This yields *HCE*=1.675 or 1,675 USD respectively, which is in the same range as the critical value estimated in Zweifel et al. (2005).

that the countries in our sample operate on the flat-of-the curve.

However, this traditional view on the production of health may well neglect the impact of HCE on the uncertainty surrounding life expectancy. This is addressed in Section 2.5.2.

2.5.2 Variability of age at death as the dependent variable

Table 2.3 presents (double-log) RE estimation results for sd and sd_{mode}, the two indicators of variability of age at death emphasized here. For a comparison with Table 2.2, elasticities of LEF_{60} evaluated at the means are provided in the column entitled Table 2.2. Three things are noteworthy. First, the same inputs that were found to increase (decrease) the expected value are estimated to decrease (increase) the variability of life expectancy. Second, whereas GDP is more effective than HCE in increasing the expected value (see fifth column), it tends to be less effective in reducing its variability of longevity. Third, HCE exhibits decreasing returns also as an instrument for controlling variability of health status.

Table 2.3: Determinants of variability of age at death, 1960-2005

	sd			Table 2.2	sd_{mode}		
	Coef.	z	P>z	ϵ	Coef.	z	P>z
VAD				LEF_{60}			
HCE_{-5}	-0.072	-3.09	0.002	0.028	-0.019	-0.27	0.788
HCE^2_{-5}	0.005	2.65	0.008		0.002	0.28	0.783
GDP	-0.066	-1.99	0.046	0.043	-0.006	-0.03	0.980
GDP^2	0.004	0.94	0.345		0.001	0.06	0.955
ALC_{-10}	0.049	1.29	0.198	-0.009	0.210	2.17	0.030
ALC^2_{-10}	0.017	1.87	0.061		-0.046	-1.97	0.049
constant	3.435	11.47	0.000		1.321	1.43	0.154
rho_{ar}	0.786				0.237		
Wald χ^2	1,103				280.46		
Prob>χ^2	0.000				0.000		
R-squared	0.6390				0.3407		
Observations	631				631		

Turning to the detailed estimation results for the overall standard deviation in Table 2.3, we find HCE_{-5} to be significant and with the predicted negative

sign. Evaluated at the mean values, a 10 percent increase of HCE five years earlier is estimated to reduce the current standard deviation of age at death by $10 \cdot (-0.072 + 2 \cdot 0.005 \cdot ln\overline{HCE}_{-5}) = 0.42$ percent. The effect of non-medical inputs is in the same range, with an increase of GDP by 10 percent associated with a decrease of variability by 0.66 percent (negelecting the insignificant squared term). As predicted, an unhealthy lifestyle proxied by ALC_{-10} seems to weaken control over health status. An earlier increase of alcohol consumption by 10 percent increases VAD by an estimated 0.49 percent (again neglecting the squared term).

For the standard deviation above the mode, the results are quite different. Only alcohol consumption is significant at the 5 percent level, with a similar estimated effect. A 10 percent increase 10 years earlier goes along with a $10 \cdot (0.210 - 2 \cdot 0.046 \cdot ln\overline{ALC}_{-10}) = 0.26$ percent increase of the standard deviation above the mode. Especially at older ages, unhealthy lifestyle in the past seems to induce lack a of control over health status. Still, the insignificant coefficients pertaining to HCE come as a surprise because according to e.g. Miller and Frech (2000), health status of the elderly (measured by their remaining life expectancy) appears to have strongly benefited from pharmaceutical innovation in particular. The apparant contradiction may be resolved by referring back to Figure 2.3. There, it appears that HCE may have influenced variability of age at death among the elderly only in recent years, possibly due to medical progress for the treatment of old-age diseases (e.g. circulatory and respiratory diseases and cancers). The graph suggests reestimation of the model for the time period between 1983 to 2005. Results are presented in Table 2.4 below.

Now, HCE_{-5} turns out to be significant at the 5 percent level, with a 10 percent increase serving to reduce variability of age at death by an estimated 0.56 percent. Almost the same magnitude is found for GDP. However, ALC_{-10} is found to be insignificant. Also note that the estimated coefficients pertaining to HCE and GDP cannot be distinguished from those for the overall sd in Table 2.2. Therefore, as to the research question Q2 stated in Section 2.2, we can conclude that both medical and non-medical inputs contribute to the observed reduction in VAD, and to a comparable extent. As to research question Q3, the answer depends on the period of observation. For the period as a whole (1960 to 2005), the elderly seem to differ in that VAD above the modal year cannot be related to either HCE or GDP. However, from the mid-1980s on, these two variables have effects that are comparable to those on the general population.

Table 2.4: Determinants of variability of age at death above the modal year (sd_{mode}), 1983-2005

Explanatory variable	Coef.	z	P>z
HCE_{-5}	-0.056	-2.49	0.013
HCE^2_{-5}	0.005	0.47	0.639
GDP	-0.058	-2.97	0.001
GDP^2	0.002	0.15	0.881
ALC_{-10}	0.061	0.28	0.776
ALC^2_{-10}	-0.017	-0.36	0.719
constant	1.774	5.52	0.000
rho_{ar}	0.217		
Wald χ^2	2,284		
Prob>χ^2	0.000		
R-squared	0.2921		
Observations	430		

2.6 Is flat-of-the-curve medicine wasteful?

In Section 2.5.1, evidence was presented to the effect that many OECD countries presently are characterized by flat-of-the-curve medicine if the marginal contribution of HCE to remaining life expectancy is accepted as the criterion. However, according to the conclusion of Section 2.5.2, HCE does contribute to reduced uncertainty with regard to health status. In the present context, this effect is valued using the risk premium an individual would be willing to pay for reducing the risk of premature death indicated by the variability of age at death (VAD).

Theoretically, the risk premium can be derived from the following equality condition that makes an individual indifferent between the certain health status A after deduction of the premium $\rho(\tilde{X})$ and the risky health profile B of Figure 2.1. In Eq. (2.6) below, $u[H_A]$ denotes the certain utility associated with certain health A (in money equivalent) and EU, expected utility associated with risky health B, which is composed of H_A and a small variation $\tilde{X}$ of health status,

$$u[H_A - \rho(\tilde{X})] = EU[H_A + \tilde{X}]. \tag{2.6}$$

Applying Taylor approximations to both sides and solving for $\rho(\tilde{X})$, one obtains the Arrow-Pratt formula in terms of health rather than wealth (see Arrow, 1970, ch. 3)

$$\rho(\tilde{X}) = \frac{1}{2}\sigma_x^2 \cdot R_A, \tag{2.7}$$

with $R_A := -\frac{u''[H_A]}{u'[H_A]}$ defining the coefficient of absolute risk aversion. The risk premium therefore is given by the product of (one half of) the variance of health status, σ_x^2, and the coefficient of absolute risk aversion. Now, risk aversion with regard to a variation in health may well differ from risk aversion with regard to wealth. Therefore, it is important to use an estimate that has a close connection to health. The one by Friedman (1974) qualifies because it is derived from the choice of health insurance. His value of R_A is $3 \cdot 10^{-3}$; in the interest of a conservative estimate of the risk premium, we use a value of R_A equal to 10^{-4}. The next step is to express the variance of length of life (as an indicator of health), σ_x^2, in monetary units. This will be done for the two countries that devote very high per-capita amounts to health care and therefore likely constitute two extreme cases of flat-of-the curve medicine, the United States and Switzerland. The United States Environmental Protection Agency (EPA, 2000) has been using a value of USD 6.3 mn. per statistical life in its cost-benefit analyses since 1999. Baranzini and Luzzi (2001) estimate an average value of CHF 12.5 mn. for Switzerland based on labor market data as of 1995. Taking the base year 2000 and an interest rate of 3 percent, this amounts to a value of a statistical life of USD 6.5 mn. for the United States and USD 8.7 mn. for Switzerland (with and an exchange rate of 1CHF=0.6USD).

Given that the two estimates above relate to statistical lives and hence are the result of a linear extrapolation of small changes in survival porbabilities, it is also admissible to interpret them as linear extrapolations of a change of one year of life expectancy. With average life expectancies of 73.8 years (United States) and 76.2 (Switzerland) respectively, one statistical year of life is worth USD 87,927 in the United States and USD 114,102 in Switzerland. Based on a utility-theoretic model of preferences over length of life, Edwards (2008) predicts that an individual would be willing to trade one-half a year of additional life expectancy against a reduction of uncertainty by one standard deviation. Therefore, a change of one sd in age at death can be valued at some USD 43,963 (United States) and USD 57,051 (Switzerland), respectively. According to the estimation results in Section 5.2, a 10 percent increase of HCE is estimated to reduce the sd by 0.42 percent, i.e. from 18.01 to 17.93 years in the US

and from 16.24 to 16.17 in Switzerland.

Inserting these estimates into Eq. (2.7) we obtain the following willingness-to-pay (WTP) values for such a reduction:

$$\frac{1}{2}((18.01)^2 - (17.93)^2)(4.40 \cdot 10^4)^2 \cdot 10^{-4} = 2.78 \cdot 10^5 \text{ (United States)} \tag{2.8}$$

$$\frac{1}{2}((16.24)^2 - (16.17)^2)(5.71 \cdot 10^4)^2 \cdot 10^{-4} = 3.25 \cdot 10^5 \text{ (Switzerland)} \tag{2.9}$$

Distributed over 73.8 years this becomes a WTP value for the United States of USD 3,771 and USD 4,261 for Switzerland.

The last step concerns the marginal cost. In 2000, the United States spent USD 4,704 per capita on health care and Switzerland, USD 3,529. Hence, 10 percent more HCE amounts to USD 470 and USD 353, respectively. The comparison with the estimates in Eqs. (2.8) and (2.9) clearly shows that in both countries, WTP for increased certainty with regard to age at death exceeds their marginal cost in terms of HCE. Therefore, one can answer Q4 by concluding that even if HCE should not prolong life anymore, it may be worth its cost as "real insurance" reducing the variability of health status.

2.7 Conclusion

This study addresses an issue that has been overlooked in the production of health literature with its emphasis on flat-of-the-curve medicine. For risk-averse individuals, not only the level of health but also its variability is important. However, improved control over health status is reflected in an increased rectangularization of the survival curve, indicating a reduced variability of age at death (VAD). Since this rectangularization can indeed be observed in OECD countries, this raises four research questions. Are the countries of our sample characterized by flat-of-the-curve medicine (Q1)? Do medical or non-medical factors contribute more to reducing VAD (Q2)? Do these effects differ among the elderly, where rectangularization has been prominent (Q3)? Is flat-of-the-curve medicine wasteful (Q4)? The standard deviation (sd) of age at death serves as an indicator of overall uncertainty concerning health status and the sd above the mode (where the number of deaths in adulthood reaches its maximum) as an indicator of uncertainty surrounding health status among the elderly. Between 1960 and 2005 both

measures decreased for the 24 OECD countries sampled, pointing to reduced VAD. However, sd above the mode began to fall in the early 1980s only. These two indicators are related to HCE as a proxy of medical inputs to the production of health and to GDP and alcohol consumption as a proxy of non-medical ones. Based on a specification that takes account of hidden heterogeneity through random effects, the four research questions can be answered as follows.

Q1: According to our estimates the critical value of HCE beyond which its marginal effect ceases to be positive can be put at USD 1,675. With a mean value of USD 3,436 as of 2005, the OECD countries of our sample are on average well within the flat-of-the-curve range.

Q2: The reduction of VAD (indicating better control over health status) is importantly due to both, HCE and GDP.

Q3: Significant effects of HCE and GDP on VAD among the elderly are found for the time period between 1983 to 2005 only, of a magnitude comparable to Q2.

Q4: Comparing the marginal cost in terms of HCE with the willingness-to-pay values for the United States and Switzerland, we find that the benefits in terms of reduced VAD exceed the extra cost. Therefore, flat-of-the-curve medicine may be worthwhile as "real insurance" serving to reduce unvertainty of health status.

However, several limitations of this study need to be pointed out. First, variability of health status as experienced by individuals is only crudely measured by cross-sectional measures such as the standard deviation of age at death. Tracking individual's health status over time would be preferable, but availability of panel data would restrict the analysis to a few countries only. Second, medical and especially non-medical inputs to the production of health are not very well captured by HCE and GDP and alcohol consumption per capita, respectively. Unfortunately, measures of education and other indicators of lifestyle do not date back sufficiently far for many OECD countries. Third, we used a coefficient of absolute risk aversion derived from U.S. data. Its value likely differs between countries.

However, the findings on the whole do suggest that variability of health status can be influenced. This has important implications. First, reduced uncertainty about age at death likely has been modifying the decisions especially of older individuals concerning savings, consumptions, and the purchase of life and long-term care insurance. Quite

generally, it helps risk-averse individuals to optimize lifetime consumption, permitting them to reduce precautionary saving (see Palumbo, 1999 and Levhari and Mirman, 1977). Second, knowing the extent and determinants of variability of health status enables insurers and reinsurers to calculate more accurate values of the financial risk they are exposed to and expected to face in the future in different countries. Finally, our study suggests complementing the economic evaluation of medical interventions (such as cost-utility or cost-effectiveness analysis) with possible reductions in the uncertainty of outcomes.

Acknowledgments

The authors would like to thank Philippe Widmer, Frank Lichtenberg, and participants of the 42^{nd} Annual Meeting of the Canadian Economic Association and of the 7^{th} European Conference on Health Economics for their helpful comments.

Appendix

In general terms, the Gini coefficient is defined as the area between the diagonal and the Lorenz curve, divided by the whole area below the diagonal (see Figure 2.4). In applications to income, the Lorenz curve represents the cumulative income share as a function of the cumulative population share (Lorenz, 1970). Following Hanada (1983), the Gini coefficient can be applied to life tables as follows. Let x be years lived rather than income. In order to measure the number of years lived, the person's death must be observed. Therefore the density function of x is redefined as

$$f_x = \frac{d_x}{l_0}, \tag{2.10}$$

with d_x denoting the number of deaths at age x and l_0 the number of survivors at year 0 (standardized to 100,000). The cumulative distribution function can be written as

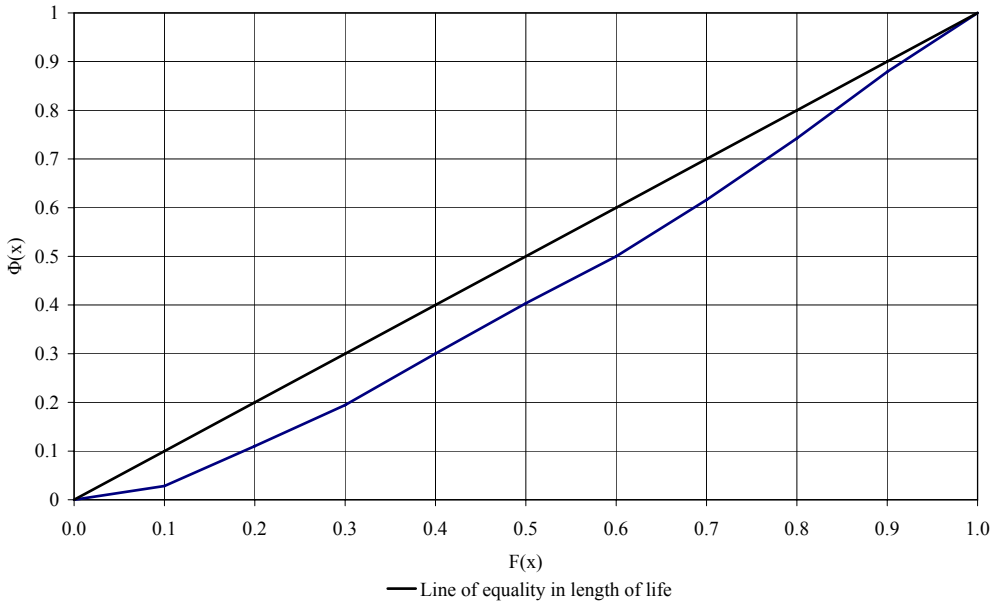
$$F_x = \sum_{x=0}^{n-1} f_x. \tag{2.11}$$

It defines the horizontal axis of Figure 2.4, with n denoting the oldest age in the life table. The share of the total amount of years lived by the share F_x of the population is

$$\Phi_x = \sum_{x=0}^{n-1} \left(\frac{d_x x}{\sum_{x=0}^{n-1} d_x x} \right), \tag{2.12}$$

representing the vertical axis of Figure 2.4. The Lorenz curve is defined over $[0, 1]$, the range of F_x. In a situation of perfect equality, the share of the population F_x coincides with its share in the total of life years lived, Φ_x. Therefore, the Lorenz curve runs diagonal in this case, from point $(0, 0)$ to $(1, 1)$. The higher the variability in years lived across a population, the greater the divergence between the diagonal and the Lorenz curve. Noting that the total area below the diagonal is 0.5 and integrating the

Figure 2.4: Lorenz curve for length of life, U.S. males (1960)



Source: Human Mortality Database (2008).

areas stepwise[7], one obtains the Gini coefficient by using Eqs. (2.11) and (2.12),

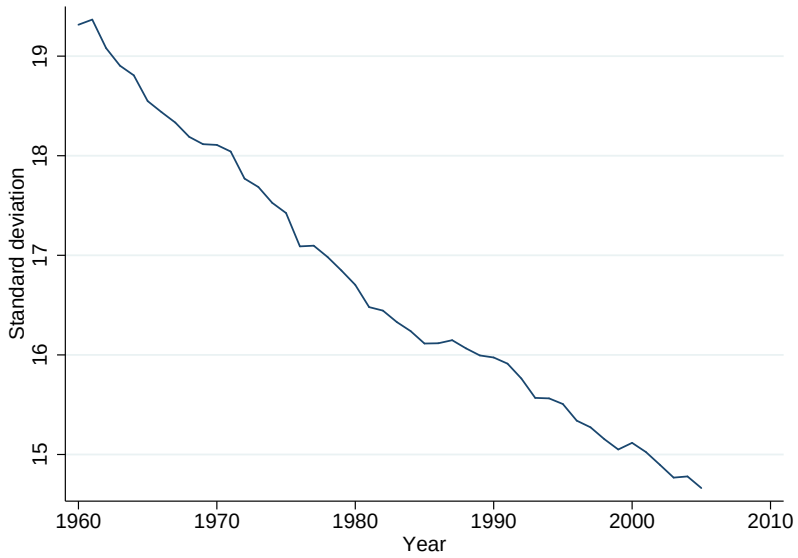
$$\begin{aligned} G &= \frac{\frac{1}{2} \sum_{x=0}^{n-1} (F_x - F_{x+1})(F_x - \Phi_x + F_{x+1} - \Phi_{x+1})}{\frac{1}{2}}, \\ &= \sum_{x=0}^{n-1} (F_{x+1} - F_x)(F_x - \Phi_x + F_{x+1} - \Phi_{x+1}). \end{aligned} \tag{2.13}$$

The Gini coefficient varies between 0 (perfect equality and hence minimum uncertainty) and 1 (perfect inequality and hence maximum uncertainty). It is equal to 0 if all

[7]The area between the diagonal and the Lorenz curve can be divided into trapezoids.

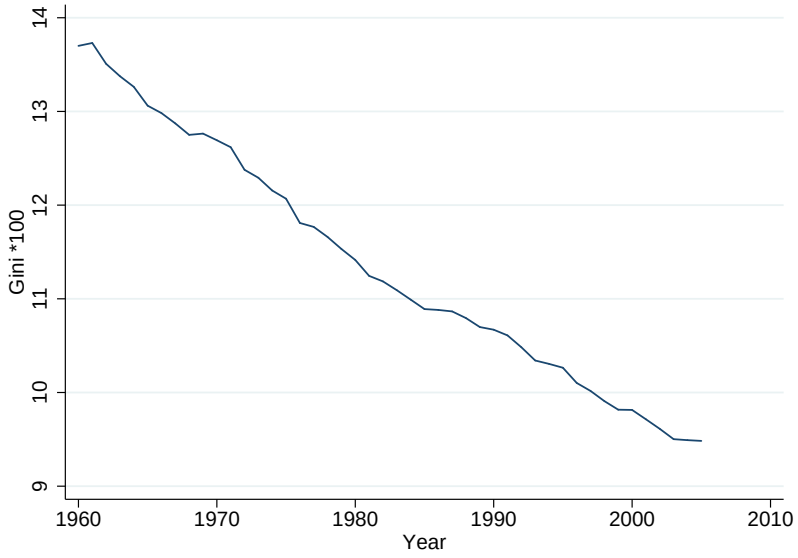
individuals of the hypothetical birth cohort die at the same age (live to the same age, respectively) and equal to 1 if everyone dies at age 0 while one individual dies at the maximum age.

Figure 2.5: Standard deviation of age at death averaged over 24 OECD countries, 1960 to 2005



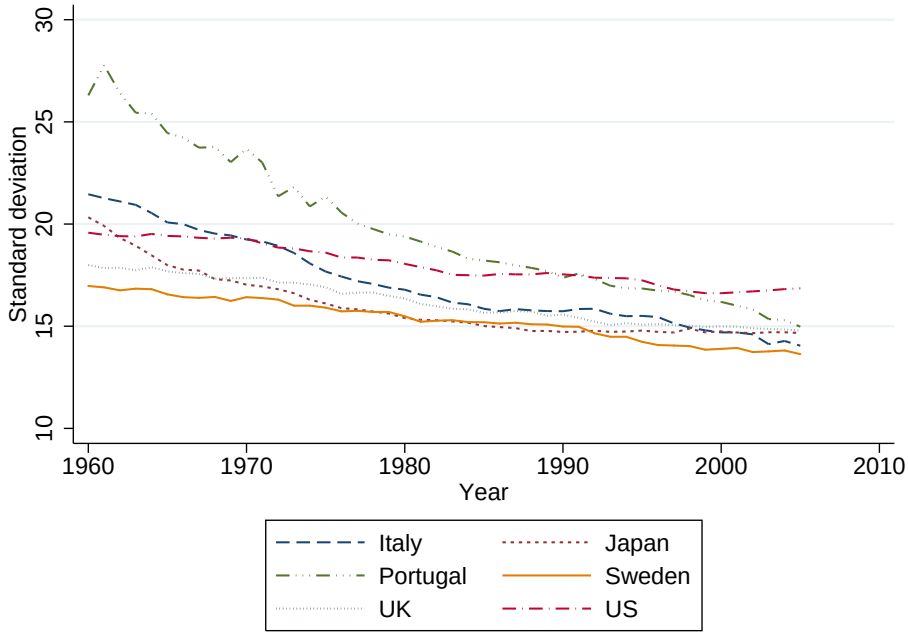
Source: Human Mortality Database (2008).

Figure 2.6: Gini coefficient averaged over 24 OECD countries, 1960 to 2005



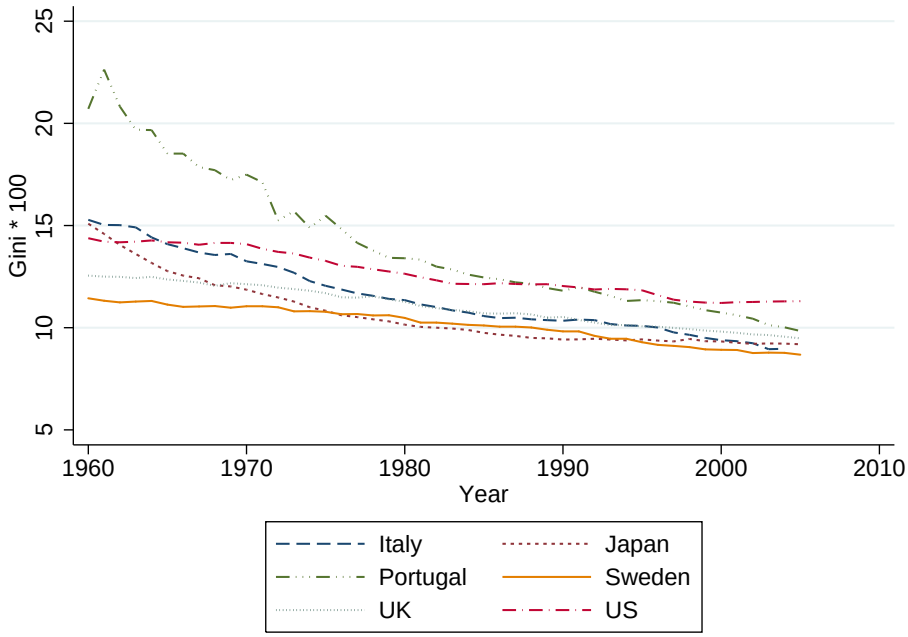
Source: Human Mortality Database (2008).

Figure 2.7: Standard deviation for selected countries, 1960 to 2005



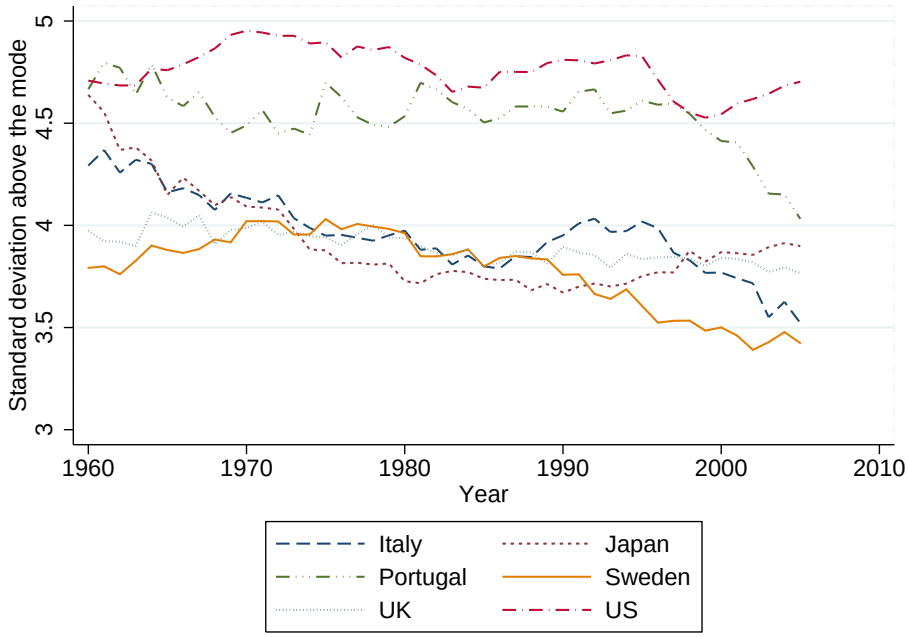
Source: Human Mortality Database (2008).

Figure 2.8: Gini coefficients for selected countries, 1960 to 2005



Source: Human Mortality Database (2008).

Figure 2.9: Standard deviation above the mode for selected countries, 1960 to 2005



Source: Human Mortality Database (2008).

Chapter 3

Impact of Specialization on Health Outcomes - Evidence from U.S. Cancer Data

Johannes Schoder and Frank Lichtenberg

Revise and Resubmit

Health Services Research

3 Impact of Specialization on Health Outcomes - Evidence from U.S. Cancer Data

3.1 Introduction and motivation

According to Lichtenberg (2007c) life expectancy at birth in 2004 differed by up to 8 percent between U.S. states. Regional differences in survival rates for specific diseases are even higher. E.g., 5-year survival rates for breast cancer vary by more than 10 percent across different U.S. regions. In this study we argue that economies of scale in the production of health explain part of the differences in cancer survival rates.

The existence of economies of scale in the production of health are quite plausible. Lower production costs as output increases may be realized because expensive medical equipment can be used more frequently and because of specialization in combination with more disease-specific knowledge and experience. In addition, these factors may also contribute to improved health outcomes. Moreover, the presence of economies of scale facilitates concentration processes which in turn are the source of productivity and knowledge spillovers further contributing to improved health outcomes (Krugman, 1991). Knowledge spillovers in medical care have been first analyzed by Coleman et al. (1957), who find that physicians integrated in the community of their colleagues are the first to adopt new drugs. Recent studies suggest that hospitals surrounded by higher quality hospitals tend to improve in quality (Baicker and Chandra, 2010). In all, one may expect that regions where more health is "produced" than on average realize above average outcomes. Using data from the National Cancer Institute's Surveillance Epidemiology End Result[1] (SEER) we analyze whether areas with relativly more of the same type of cancer indeed exhibit above average survival rates.

[1]This database pools together information from all U.S. cancer registries on where and how many people are diagnosed with cancer and their survival time.

The study contributes to the literature of specialization as follows. So far it is widely recognized that physicians or hospitals improve their productivity when specialized on a narrow range of tasks. The positive relationship between volume and outcome is well established (Gillis and Hole, 1996, Birkmeyer et al., 2003, and Allgood and Bachman, 2006). Based on these results a lot of studies recommend to centralize health care provision in specialized units. E.g. in Maryland, health authorities require hospitals to perform more than 200 open heart surgeries per year in order to keep their program in good standing (Sfekas, 2009). However, especially for cancer treatment an analysis on the level of a physician or hospital is too narrow for two reasons. First, physicians not only work for one hospital, they often have operating privileges at multiple hospitals and interact (socially and professionally) with other doctors. Second, when being treated for cancer it is very common for patients to be diagnosed in one facility, receive surgery in a second facility, have radiation therapy in yet another facility, and be treated on an outpatient basis with chemotherapy.[2] Through such interactions, knowledge spillovers are expected to reach beyond the boundary of a hospital to affect health outcomes of all patients in a region (Chandra and Staiger, 2007). Thus, possible specialization effects may remain undetected.

The paper is structured as follows. Section 3.2 starts with a simple model on cancer survival. Section 3.3 is devoted to the data source and the empirical estimation strategy. The estimation results are presented in Section 3.4. Section 3.5 concludes.

3.2 A simple model of cancer survival

The literature distinguishes three different types of survival rates (see e.g. Horner et al., 2008). The observed survival rate (S) is the probability of surviving all causes of death for a specified time after diagnosis of cancer. It considers deaths from all causes, cancer or otherwise. The expected survival rate (E) is the survival rate of a comparable set of people that do not have cancer. In turn, the relative survival rate (R) is defined as the ratio of the proportion of observed survivors (all causes of death) in a cohort of cancer patients to the proportion of expected survivors in a comparable cohort of cancer-free individuals, S/E.

[2] E.g. the successful treatment of esophageal cancer needs an experienced team of inpatient specialists as well as experienced outpatient oncologists for the follow-up treatment (Hoelscher, 2001).

The following model is based on Lichtenberg (2007b) and explains the survival rate as a function of different input factors. However, instead of using individuals as the unit of observation we use regions. We assume that the relative survival rate depends on the treatment quality and the disease progression at time of diagnosis:

$$R_{ir} = S_{ir}/E_{ir} = f_1(Q_{ir}, P_{ir}) \tag{3.1}$$

or

$$S_{ir} = f_2(E_{ir}, Q_{ir}, P_{ir}) \tag{3.2}$$

with,

- R_{ir}: The relative survival rate for cancer type i in region r
- S_{ir}: The observed survival rate for cancer type i in region r
- E_{ir}: The expected survival rate of the control group
- Q_{ir}: Treatment quality for cancer type i in region r
- P_{ir}: Disease progression of cancer type i in region r

The observed survival rate is hypothesized to be an increasing function of the quality of treatment ($\frac{\partial f_2(\cdot)}{\partial Q_{ir}} > 0$) and the expected survival rate ($\frac{\partial f_2(\cdot)}{\partial E_{ir}} > 0$) and a decreasing function of disease progression at time of diagnosis ($\frac{\partial f_2(\cdot)}{\partial P_{ir}} < 0$).

Quality in health care is hard to define and hard to quantify. Lichtenberg (2007a) uses treatment vintage to measure treatment quality.[3] However, treatment quality is also a function of knowledge and the extent of specialization. According to Birkmeyer et al. (2002) low volume hospitals with less than 3 pancreatic resections per year report a 11 percent higher mortality rate than high volume hospitals with more than 16 cases. In a follow-up study similar results are found on the level of physicians (Birkmeyer et al., 2003). Thus, we hypothesize that treatment quality is an increasing function of specialization. In turn, specialization is measured by N_{ir}, the number of cancer type i in region r:

$$Q_{ir} = f_3(N_{ir}) \tag{3.3}$$

Substituting Eq. (3.3) into Eq. (3.2):

[3]The vintage of a treatment is the year in which the treatment was first used.

$$S_{ir} = f_2(E_{ir}, P_{ir}, N_{ir}). \tag{3.4}$$

Our primary objective is to estimate the effect of specialization (N_{ir}) on survival (S_{ir}). Thus our testable hypothesis is $\frac{\partial f_2(\cdot)}{\partial N_{ir}} > 0$.

The observed survival rate could depend on factors other than specialization and disease progression. It likely depends on the medical infrastructure, the availability of new chemotherapy drugs (treatment vintage), and patient characteristics such as income. However, since we will include region and disease dummies we are able to control for these unobserved characteristics.

3.3 Econometric model and Data

The model in Section 3.2 includes the expected survival rate.[4] However, the SEER data does not allow calculation of expected survival rates according to the county level. Therefore additional population characteristics such as age, race, and sex are included for the county estimation.[5] Based on Section 3.2 we estimate the following model.

$$\begin{aligned} S_{ir} &= \beta_1 N_{ir} + \beta_2 Exp_r + \beta_3 LOC_{ir} + \beta_4 DIST_{ir} + \beta_5 SURG_{ir} + \beta_6 RAD_{ir} \\ &\quad + \beta_7 AGE_{ir} + \beta_8 WHITE_{ir} + \beta_9 MALE_{ir} + \epsilon_{ir} \end{aligned} \tag{3.5}$$

where

- r: Place of diagnosis (cancer registry or country); i: Cancer type.
- S_{ir}: Observed survival rate for cancer type i and region r. Survival rates will be calculated for five different time intervals.
- N_{ir}: Number of people diagnosed with cancer type i in region r. According to our hypothesis we should expect a higher survival rate for cancer type i in registries with more diagnosed cancers of type i. This effect should be independent from region and disease heterogeneity.

[4] Controlling for the expected survival rate leaves us with the survival rate of people having cancer.

[5] However, according to Lichtenberg (2007b) excluding the expected survival rate poses little risk of biasing the right hand side variables since we control for the mean age of people diagnosed.

- Exp_i: Expected survival rate of the control group. This variable controls for all factors that influence survival in general (e.g. gender, age or race) but that are not related to cancer treatment. We should observe a positive impact.

The next four variables indicate by how much the cancer has spread and are used as a measure of mean progression of disease.

- LOC_{ir}: Share of cancer stage I or II for cancer at site i in region r in 2002. In stage I and II cancers are localized to one part of the body. The higher the share of cancers in stage I or II, the higher the survival rate is likely to be.

- $DIST_{ir2002}$: Share of cancer stage IV for cancer site at i in region r in 2002. In stage IV cancers have often metastasized, spread to other organs, or spread throughout the whole body. The higher the share of cancers in stage IV, the lower the survival rate is likely to be.

- $SURG_{ir2002}$: Share of patients receiving surgery for cancer at site i in region r in 2002. When a surgery is performed as a primary treatment, chances are high for cure, especially if the cancer is localized and has not spread. However, a surgery can also be performed in order to remove as much as possible of the tumor in order to make chemotherapy or radiation more effective or to just improve the quality of life. Unfortunately we only know whether a surgery was carried out or not.

- RAD_{ir}: Share of patients receiving radiation for cancer type i in region r. Radiotherapy may be used as therapeutic treatment where the therapy has survival benefit or as palliative treatment where cure is not possible anymore. Therefore the effect of RAD is unclear.

Since the expected survival rate is not available on the county level we include the following individual characteristics in the county level estimation:

- AGE_{ir}: The mean age of people diagnosed with cancer type i. The literature shows that among adults, relative survival decreases with increasing age at diagnosis for almost every cancer (Cancer Research UK, 2006). Therefore we should expect lower survival rates for counties where the mean age of people diagnosed with cancer is higher. However we have to be careful with the interpretation of the coefficient of AGE since it also picks up the effect of earlier detection.

- $WHITE_{ir}$: Share of white patients diagnosed with cancer type i. Several studies find higher survival rates for whites compared to blacks (Pulido et al. (2009), Coleman et al. (2008)). Since the fixed effects model already controls for disease-invariant characteristics the variable can also be interpreted as life style differences between whites and blacks. We hypothesize that counties with a higher share of whites exhibit higher survival rates.

- $MALE_{ir}$: Share of male patients diagnosed with cancer type i. There are certain cancer types that are more prevalent among women (e.g. breast cancer) or among men (e.g. prostate cancer). Since we include cancer type fixed effects we control for these differences. However, the literature on the production of health finds consistently higher survival rates for females than for males (see e.g. Miller and Frech (2000)). Therefore one can think that there still remain differences in the survival rates due to gender even after controlling for different cancer types.

- ϵ_{ir}: Error term

Data are obtained from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program. It contains information from population-based cancer registries[6] covering approximately 26 percent of the US population. The registries included are San Francisco-Oakland (since 1973), Connecticut (1973), Detroit (1973), Hawaii (1973), Iowa (1973), New Mexico (1973), Utah (1973), Seattle (1974), Atlanta (1975), Alaska (1992), San-Jose Monterey (1992), Los Angeles (1992), Rural Georgia (1992), remaining California (2000), Kentucky (2000), Louisiana (2000-2004), and New Jersey (2000). The variables included are patient demographics, primary tumor site, tumor morphology and stage at diagnosis, first course of treatment, and follow-up for vital status. The variable of interest in our study is the survival rate. Survival rates may be calculated for different time intervals. To assess treatment effects for cancer the literature usually refers to 5-year survival rates. In addition we also estimate models for the 1-, 2-, 3-, and 4-year survival rate.

For the purpose of our analyis we group each individual-based cancer record according to region and aggregated cancer site. SEER provides two different geographical identifiers (unfortunately not the hospital). We will use both, the registry and the

[6]Cancer registries are a systematic collection of data about cancer and tumor diseases. The geographic area of one SEER registry corresponds to approximately one U.S. state, except for the Californian registries, Seattle, rural Georgia, Atlanta, and Detroit.

county identifier, for the estimation. Table 3.1 contains the summary statistics for the registry level according to aggregated sites.[7] In total we are left with 272 observations at the registry level. Survival rates are weighted by the number of people diagnosed. Chances of surviving cancer one year, three years, and five years after diagnosis are 78 percent, 65 percent, and 58 percent respectively. The differences between survival rates across registries are up to 23 percentage points. Table 3.2 shows the number

Table 3.1: Summary statistics (registry level, aggregated sites) 1998 to 2002

Variable	Mean	sd	Min	Max
1-year survival	0.77	0.19	0.21	1
2-year survival	0.69	0.23	0.16	1
3-year survival	0.64	0.24	0.16	1
4-year survival	0.60	0.25	0.14	1
5-year survival	0.57	0.24	0.12	1
1-year exp. survival	0.98	0.01	0.96	1
2-year exp. survival	0.95	0.02	0.91	1
3-year exp. survival	0.93	0.03	0.87	1
4-year exp. survival	0.90	0.04	0.82	1
5-year exp. survival	0.88	0.05	0.78	1
N	4,004	5,935	2	33,424
lnN	7.06	1.95	0.69	10.42
AGE	58.82	8.88	27.7	71.54
LOC	0.49	0.34	0	1
DIST	0.23	0.31	0	1
SURG	0.61	0.30	0	1
RAD	0.29	0.18	0	0.67
WHITE	0.77	0.25	0	1
MALE	0.51	0.25	0	1
Observations	272			

Note: The survival rates are weighted by N_{ir}.

of people diagnosed according to SEER registry and aggregated cancer site. Ideally we would have information on the place where the diagnosed patient receives its treatment, however, this is not available in the SEER data. There are two low volume registries, viz. Alaska and Rural Georgia with 260 and 509 cases respectively and three high volume registries with above 20,000 cases, viz. Los Angeles (29,232), New Jersey (38,172), and greater California (63,147). The cancers less common are at site eye and orbit (528) and bones and joint (659). In contrast, most common cancers

[7]In this study we use the 16 aggregated (to broad sites, following the National Cancer Institute) sites based on the international classification of diseases for oncology, 3rd edition.

Table 3.2: Number of diagnoses according to registry and aggregated site, 2002

Registry	0	1	2	3	4	5	6	7	8
San Francisco	371	2,891	1,924	34	101	582	2,677	920	2,843
Conneticut	304	3,020	2,216	33	104	649	2,395	901	2,809
Metropolitan Detroit	390	3,081	2,632	33	107	674	2,415	1,018	3,723
Hawaii	136	1,088	582	11	33	190	779	293	677
Iowa	289	2,480	1,802	22	66	541	1,911	744	2,116
New Mexico	143	1,111	770	16	45	293	960	391	1,223
Seattle Puget-Sound	370	2,696	2,188	33	80	842	2,773	947	2,911
Utah	129	978	424	16	41	351	891	367	1,425
Metropolitan Atlanta	233	1,474	1,170	24	78	537	1520	515	1,778
San Jose Monterey	166	1,374	831	16	61	316	1,222	434	1,434
Los Angeles	653	5,864	3,393	83	201	1,150	4,804	2,019	5,432
Alaska	5	86	47	3	3	1	41	7	27
Rural Georgia	19	97	93	4	2	14	76	27	91
$GreaterCalifornia$[1]	1,488	11,372	8,664	166	444	3,425	10,218	3,656	11,417
Kentucky	424	3,200	3,764	44	104	735	2,454	974	2,749
Louisiana	442	3,516	3,012	41	114	477	2,472	933	3,372
New Jersey	726	7,180	5,133	80	240	1,622	5,412	2,391	7,837
total	6,288	51,508	38,645	659	1,824	12,399	43,020	16,537	51,864

Registry	9	10	11	12	13	14	15		total
San Francisco	923	29	227	321	797	168	376		15,184
Conneticut	1,160	25	219	319	759	197	386		15,496
Metropolitan Detroit	1,201	29	250	329	794	238	452		17,366
Hawaii	250	1	47	129	194	45	101		4,556
Iowa	943	27	201	262	648	147	429		12,628
New Mexico	411	8	92	211	264	88	205		6,231
Seattle Puget-Sound	1,099	35	268	362	842	180	441		16,067
Utah	326	14	128	235	329	70	211		5,935
Metropolitan Atlanta	478	15	149	253	438	117	206		8,985
San Jose Monterey	442	19	114	163	415	89	211		7,307
Los Angeles	1,770	66	447	747	1,451	339	813		29,232
Alaska	11	0	4	7	6	4	8		260
Rural Georgia	36	1	3	8	21	9	8		509
$GreaterCalifornia$[1]	4,246	131	1,053	1,277	3,081	721	1,788		63,147
Kentucky	1,235	41	248	284	837	217	414		17,724
Louisiana	1,243	23	219	335	757	261	457		17,674
New Jersey	2,782	64	506	936	1,806	472	985		38,172
total	18,556	528	4,175	6,178	13,439	3,362	7,491		27,6473

Note: 0: Oral cavity, 1: Digestive System, 2: Respiratory System, 3: Bones and Joints, 4: Soft tissue incl. heart, 5: Skin excl. basal and squamous, 6: Breast, 7: Female genital system, 8: Male genital system, 9: Urinary system, 10: Eye and orbit system, 11: Brain, 12: Endocrine System, 13: Lympoma, 14: Myeloma, 15: Leukemia. [1]: Excluding San Francisco, Los Angeles, and San Jose Monterey.

are at site male genital system (51,864), digestive system (51,508), and breast (43,020).

The aim of our analysis is to estimate the relationship between S_{ir}, the proportion survived, and the exogenous factors, X. As a first attempt, we will formulate the model as a linear logistic regression of S_{ir} on X, that is we will take the logit of S_{ir} and represent the response curve as a straight line:

$$ln(\frac{S_{ir}}{1 - Sir}) = \gamma_0 + \gamma_1 X \tag{3.6}$$

However, since we grouped the data (according to region and aggregated site) and these groups differ in terms of size it is not possible to use a standard logistic regression to fit the model. Thus, Eq. (3.5) is estimated using a GLM model with a logit link function and a binomial distribution function where the denominator is the number of diagnoses (see Stata, 2007). To control for disease and regional specific characteristics we include fixed effects for cancer site i and region r. A significant coefficient for the variable N_{ir} would then imply that the ratio of the odds of surviving from cancer at site B to the odds of surviving from cancer at site A is positively correlated, across regions, with the ratio of the number of patients diagnosed with cancer at site B to the number of patients diagnosed with cancer at site A (ceteris paribus, generalized to i cancer sites).

3.4 Estimation results

Table 3.3 contains the estimation results for the five different survival intervals. The standard errors are in parentheses and clustered according to each registry since observations within registries are possibly correlated (see Bertrand et al., 2002). Furthermore, since the estimation is based on grouped data (defined by cancer type and registry) and these groups differ in terms of size we weight the equation by the number of patients diagnosed. Based on the Akaike and Bayes information criteria a model specification with the logarithmic of the number of diagnosis is preferred pointing to diminishing marginal returns of specialization (possibly due to increasing coordination costs, see e.g. Becker and Murphy, 1992). In general, most coefficients are of the expected signs. The effect of the variable of main interest, number of diagnosis, is positive and significant across all five survival intervals confirming our hypothesis that specialization leads to higher survival chances. To illustrate the effect, consider the following comparison. Focusing on cancer of the respiratory system, the registry

Table 3.3: Estimation results (registry level, aggregated sites), 1998-2002

	Coefficients for (t)-year survival rates				
	(1)	(2)	(3)	(4)	(5)
lnN	0.210***	0.186**	0.216**	0.209**	0.189**
	(0.078)	(0.090)	(0.088)	(0.088)	(0.078)
$lnEXP_t$	-5.340	0.378	0.973	1.456	1.683*
	(3.519)	(2.286)	(1.758)	(1.188)	(0.992)
LOC	0.202	-0.005	0.024	0.174	0.209
	(0.255)	(0.216)	(0.205)	(0.177)	(0.173)
DIST	-0.830	-1.013*	-1.047*	-0.975*	-1.078**
	(0.514)	(0.526)	(0.534)	(0.532)	(0.530)
SURG	0.674**	0.594**	0.539**	0.493**	0.422*
	(0.323)	(0.302)	(0.265)	(0.231)	(0.222)
RAD	0.510**	0.345	0.418*	0.405*	0.402**
	(0.246)	(0.245)	(0.229)	(0.208)	(0.201)
AIC	9.88	10.69	10.96	10.88	10.82
BIC	-523.83	-394.03	-354.15	-392.56	-416.87
LogLikelihood	-1,306	-1,416	-1,452	-1,441	-1,433
Observations	272	272	272	272	272

Note: Fixed registry and cancer type effects are included, see Appendix; robust standard errors are given in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

Los Angeles has the highest number of people diagnosed, 17,687 compared to the registry average of 9,048. Although Los Angeles has above average numbers of cancer of the respiratory system its 1-year survival rate lies 2.2 percentage points below (40.6 percent) the average (weighted) 1-year survival rate of all registries in the sample (42.8 percent). However, this is due to the lower incidence rate of 12.3 percent for Los Angeles (here defined as the share of cancer of the respiratory system to all cancers within the registry Los Angeles) against 14.1 percent for the average incidence rate of all registries. Hence, the *relative* number of cancers is crucial in determining survival chances. Now, consider a registry with only few people diagnosed with cancer of the respiratory system. There are 232 people diagnosed in the registry Alaska but the survival rate of 42 percent - although lower than the registry average - is still higher than the registry Los Angeles. Again this is due to the high incidence rate of cancer of the respiratory system of 17.9 percent, well above the registry average of 14.1 percent.

The value of the expected survival rate (remember this is the survival rate of a comparable set of people not having cancer) is only significant for the 5-year survival

rate, but has the expected positive sign in all specifications (except for the 1-year survival rate). An important determinant is the treatment variable surgery whereas the variable radiation is only significant at the 10 percent level. SURG has a positive and significant impact across all survival rates. A one unit increase in the share of surgeries almost doubles the odds of being alive after one year. However, the effect decreases for the longer survival intervals. The progression rate of cancer does not seem to be an important determinant of survival. Only metastasized cancer reduces survival chances considerably.

To test the robustness of our model we also estimated Eq. (3.5) on the county level. The 17 registries are divided into 471 different counties[8] leading to 7,093 number of observations. Since SEER does not provide the expected survival rate on the county level we include age (AGE), race (WHITE) and sex (MALE) to control for population-specific characteristics. Table 3.4 includes the estimation results. In general, as the sample size increases more coefficients tend to turn out significant. The variable number of diagnosis is again positively related to the survival rate and significant at the 1 percent level (except for the 5-year survival rate estimation). However, this time the effects are smaller in terms of magnitude. E.g. the highest effect is found for the 2-year survival rate. Here, a one unit increase in the log of number of diagnoses increases the odds of being alive after two years by more than 13 percent conditional on county and cancer type survival mean. The lower effects are likely due to the fact that specialization operates at a broader level than that of a given county, e.g. counties sometimes only have one hospital. Moreover, since patients are treated and diagnosed at different locations and counties sometimes only cover small areas, results are possibly biased downwards. As expected the variable AGE is negative and significant across all five specifications. An increase of the mean age at diagnosis reduces survival. As in the previous estimation only the DIST variable is significant and negative across all five specifications. The treatment variable SURG also plays a more important role than RAD. The higher the share of whites in a county the higher are survival chances which confirming the findings of Pulido et al. (2009) and Coleman et al. (2008). In contrast, increases of the share of males tend to decrease survival chances. To sum up, the results for the county level estimation are similar to the registry level estimation providing evidence on the robustness of the

[8]We also performed a detailed estimation for the disaggregated sites, where we have 78 different cancer types. For the most detailed estimation we are left with 26,363 observations

Table 3.4: Estimation results (county level, aggregated sites), 1998-2002

	Coefficients for (t)-year survival rates				
	(1)	(2)	(3)	(4)	(5)
lnN	0.096***	0.135***	0.134***	0.104***	0.054**
	(0.032)	(0.029)	(0.027)	(0.024)	(0.022)
AGE	-0.025***	-0.026***	-0.026***	-0.027***	-0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LOC	0.189	0.052	0.093	0.097	0.112
	(0.127)	(0.118)	(0.107)	(0.089)	(0.105)
DIST	-1.842***	-2.005***	-2.016***	-2.008***	-1.604***
	(0.150)	(0.148)	(0.137)	(0.124)	(0.151)
SURG	0.868***	0.806***	0.707***	0.602***	0.320***
	(0.088)	(0.089)	(0.078)	(0.072)	(0.064)
RAD	0.181**	0.060	0.085	0.086	0.027
	(0.084)	(0.071)	(0.066)	(0.056)	(0.063)
WHITE	0.220*	0.278***	0.233***	0.301***	0.297***
	(0.120)	(0.101)	(0.0879)	(0.0832)	(0.0724)
MALE	-0.203***	-0.315***	-0.356***	-0.347***	-0.249***
	(0.065)	(0.061)	(0.062)	(0.060)	(0.074)
AIC	4.28	4.52	4.6	4.65	4.74
BIC	-50,926	-50,687	-50,736	-50,793	-50,109
LogLikelihood	-14,667	-15,518	-15,807	-15,995	-16,306
Observations	7,093	7,093	7,093	7,093	7,093

Note: Fixed county and cancer type effects are included; robust standard errors are given in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

estimated model.

Our study is subject to some limitations. First, the assignment of a given stage to a particular cancer may change over time due to advances in diagnostic technologies. Introduction of new technology can give rise to a phenomenon known as stage migration. Stage migration occurs when diagnostic procedures change over time, resulting in an increase in the probability that a given cancer will be diagnosed in a more advanced stage.[9] The likely result would be to remove the worst survivors - those with previously undetected distant metastases - from the localized and regional categories and put them into the distant category. As a result, the stage-at-diagnosis distribution for

[9]For example, certain distant metastases that would have been undetectable a few years ago can now be diagnosed by computer tomography (CT) scan or by magnetic resonance imaging (MRI). Therefore, some patients who would have been diagnosed previously as having cancer in a localized or regional stage are now diagnosed as having cancer in a distant stage.

a cancer may become less favorable over time, but the survival rates for each stage may improve (Feinstein et al., 1985). However, since we focused on a given time period impact of the introduction of new technologies are limited. Second, cancer survival studies are often criticized that they find improved survival only because of improved earlier detection and diagnosis of cancers - caused e.g. by new screening procedures. As the proportion of cancers detected at screening increases, presumably as a result of increased screening of the population, patient survival rates will increase, because they are based on survival time after diagnosis.[10] However, since we are controlling for expected survival and cancer progression we implicitly controlled for this lead-time bias. Thus, the increased survival rates we find are mainly due to specialization gains.

3.5 Conclusion

Since specialization is often made responsible for the growth in U.S. health care expenditure, it is important to know whether more specialization is justified on the ground of improved health outcomes. So far the results are mixed, e.g. a study by Baicker and Chandra (2004) cannot find improved health outcomes for areas having a relatively higher share of specialists. By way of contrast, our study suggests that there are specialization gains on the regional level for the treatment of cancer. From a simple model of cancer survival we derived the testable hypothesis whether regional specialization in the treatment of cancer increases life expectancy. Specialization is measured by the number of diagnosed cancers. Using data from the National Cancer Institute's Surveillance Epidemiology End Result (SEER) we estimated a GLM model with a binomial distribution and logit link function. The results could not reject our hypothesis. Patients tend to survive longer in those areas where relative more cancers of the same type exist than the US average. Possibly, a higher prevalence of cancers in some regions has led to greater accumulation of disease-specific knowledge (also possibly through physician migration) which finally contributed to improved health outcomes. Moreover, the results are robust to different units of observation. A broader definition of the geographical area (here: registries) leads to higher effects of specialization on cancer survival in terms of magnitude. Similar to the study of Becker and Murphy (1992) we find decreasing marginal returns for the variable that indicates specialization. They predicted that the degree of specialization is not only limited by

[10]The interval between the time a cancer is diagnosed by a screening procedure and the time when the cancer would have been diagnosed in the absence of screening is called lead-time (Zelen, 1976).

the extent of the market but also by coordination costs (e.g. the coordination costs of organizing a whole team of cancer specialists).

The results of our study also conveys a message to individuals diagnosed with cancer. They can increase their survival chances by moving to regions that have a relativly high number of people suffering from the same disease. While this study is one of the first analyzing specialization gains in the production of health on a regional level further evidence is needed. Future research could broaden the level of analysis to different diseases or to cross-country comparisons.

Acknowledgments

The financial support by the Swiss National Science Foundation is gratefully acknowledged. The authors would like to thank Peter Zweifel and Patrick Eugster for their valuable comments.

Appendix

Table 3.5: Summary statistics (county level, aggregated sites) 1998 to 2002

Variable	Mean	Std. Dev.	Min	Max
1-year survival	0.76	0.20	0	1
2-year survival	0.68	0.23	0	1
3-year survival	0.63	0.24	0	1
4-year survival	0.57	0.24	0	1
5-year survival	0.40	0.18	0	1
N	197.70	972.16	1	36,819
lnN	3.30	1.88	0	10.51
Age	61.18	10.91	0	94
LOC	0.47	0.37	0	1
DIST	0.23	0.33	0	1
SURG	0.60	0.34	0	1
RAD	0.27	0.24	0	1
WHITE	0.90	0.18	0	1
MALE	0.51	0.31	0	1
Observations	7,093			

Note: The survival rates are weighted by N_{ic}.

Table 3.6: Summary statistics (county level, disaggregated sites) 1998 to 2002

Variable	Mean	Std. Dev.	Min	Max
1-year survival	0.76	0.24	0	1
2-year survival	0.68	0.27	0	1
3-year survival	0.63	0.27	0	1
4-year survival	0.57	0.34	0	1
5-year survival	0.40	0.20	0	1
N	53.19	404.52	1	33,165
lnN	1.96	1.66	0	10.41
Age	63.10	13.59	0	103
LOC	0.51	0.39	0	1
DIST	0.26	0.35	0	1
SURG	0.59	0.39	0	1
RAD	0.23	0.31	0	1
WHITE	0.90	0.21	0	1
MALE	0.53	0.35	0	1
Observations	26,363			

Note: The survival rates are weighted by N_{ic}.

Table 3.7: Complete estimation results (registry level, aggregated sites), 1998-2002

	Coefficients for (t)-year survival rates				
	(1)	(2)	(3)	(4)	(5)
lnN	0.210***	0.186**	0.216**	0.209**	0.189**
$lnEXP_t$	-5.340	0.378	0.973	1.456	1.683*
LOC	0.202	-0.0052	0.024	0.174	0.209
DIST	-0.830	-1.013*	-1.047*	-0.975*	-1.078**
SURG	0.674**	0.594**	0.539**	0.493**	0.422*
RAD	0.510**	0.345	0.418*	0.405*	0.402**
Connecticut	-0.019	0.004	-0.001	0.006	0.009
Metropolitan Detroit	-0.149***	-0.129***	-0.130***	-0.121***	-0.110***
Hawaii	0.282***	0.216**	0.238**	0.211**	0.183*
Iowa	-0.059***	-0.029	-0.016	-0.002	-0.001
New Mexico	0.136***	0.112	0.133	0.126	0.0939
Seattle Puget Sound	0.012	0.036	0.036	0.051*	0.055**
Utah	0.149***	0.163*	0.200**	0.191**	0.168**
Metropolitan Atlanta	0.109***	0.082	0.090	0.090	0.076
San Jose Monterey	0.276***	0.256***	0.283***	0.268***	0.251***
Los Angeles	-0.228***	-0.212***	-0.238***	-0.234***	-0.217***
Alaska	0.860***	0.682*	0.742*	0.670*	0.560*
Rural Georgia	0.483***	0.346	0.496	0.485	0.442
Greater California	-0.291***	-0.248***	-0.264***	-0.246***	-0.228***
Kentucky	-0.074***	-0.085*	-0.072	-0.069	-0.088*
Louisiana	-0.132***	-0.136***	-0.129***	-0.121***	-0.133***
New Jersey	-0.104***	-0.097*	-0.115**	-0.097*	-0.087*
Digestive system	-1.077***	-0.814***	-0.769***	-0.681***	-0.565***
Respiratory system	-1.446***	-1.427***	-1.496***	-1.465***	-1.387***
Bones and joints	1.229***	1.022***	1.092***	1.048***	1.018***
Soft tissue incl. heart	0.507***	0.497***	0.573***	0.574***	0.572***
Skin[a]	1.364***	1.246***	1.218***	1.201***	1.202***
Breast	1.106***	1.049***	0.866***	0.786***	0.753***
Female genital system	0.333***	0.391***	0.381***	0.393***	0.444***
Male genital system	1.685***	1.526***	1.372***	1.403***	1.387***
Urinary system	-0.0373	0.182	0.241*	0.275**	0.334***
Eye and Orbit system	2.334***	1.809***	1.739***	1.554***	1.408***
Brain	-1.022***	-1.188***	-1.079***	-0.904***	-0.836***
Endocrine system	1.541***	1.773***	1.838***	1.832***	1.810***
Lymphoma	0.127	0.178	0.260	0.412**	0.445***
Myeloma	1.012***	1.010**	0.930*	0.824	0.780
Leukemia	0.755***	0.940*	1.083**	1.161**	1.283**
Constant	23.570**	-2.781	-6.005	-8.396	-9.347**
AIC	9.88	10.69	10.96	10.88	10.82
BIC	-523.83	-394.03	-354.15	-392.56	-416.87
LogLikelihood	-1,306.29	-1,416.13	-1,451.9	-1,441.15	-1,433.29
Observations	272	272	272	272	272

Note: *** p<0.01, ** p<0.05, * p<0.1. [a]Excluding basal and squamous.

Table 3.8: Estimation results (county level, disaggregated sites), 1998-2002

	Coefficients for (t)-year survival rates				
	(1)	(2)	(3)	(4)	(5)
lnN	0.075***	0.088***	0.093***	0.074***	0.028**
	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)
AGE	-0.037***	-0.038***	-0.039***	-0.038***	-0.029***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
LOC	0.231***	0.113**	0.078*	0.063	0.106**
	(0.053)	(0.050)	(0.046)	(0.042)	(0.053)
DIST	-1.344***	-1.587***	-1.637***	-1.631***	-1.307***
	(0.062)	(0.061)	(0.058)	(0.056)	(0.071)
shareSurg	1.034***	0.898***	0.794***	0.693***	0.421***
	(0.068)	(0.065)	(0.059)	(0.058)	(0.048)
RAD	0.307***	0.094**	0.069*	0.047	0.029
	(0.048)	(0.040)	(0.038)	(0.034)	(0.038)
WHITE	0.171***	0.204***	0.203***	0.237***	0.200***
	(0.057)	(0.056)	(0.052)	(0.049)	(0.038)
MALE	-0.170***	-0.236***	-0.267***	-0.256***	-0.207***
	(0.032)	(0.031)	(0.03)	(0.029)	(0.036)
AIC	2.81	2.95	3	3.03	3.02
BIC	-236,446	-235,789	-235,678	-235,606	-234,415
LogLikelihood	-36,439	-38,371	-38,935	-39,318	-39,267
Observations	26,363	26,363	26,363	26,363	26,363

Note: Fixed county and cancer type effects are included; robust standard errors are given in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Chapter 4

The Contribution of Managed Care to the Performance of Health care Systems - Evidence from Three Countries

Johannes Schoder and Peter Zweifel

Published in

Swiss Journal of Economics and Statistics,
144(3), 2008, 477-493.

4 The Contribution of Managed Care to the Performance of Health care Systems - Evidence from Three Countries

4.1 Introduction

Health care expenditure (HCE) continues to increase at a faster rate than GDP in almost all industrialized countries. Governments have tried to alleviate the pressure on their budget mainly in two different ways. One has been to limit HCE by regulation, the other, to introduce competition in an attempt to increase efficiency (Cutler et al., 2000). Since the first option has not proved too successful, more and more countries are seeking ways to enhance competition, among them by fostering Managed Care (MC). By vertically integrating health insurance and health care provision, MC may improve the allocation of resources in health care while limiting HCE.

Indeed, most of the current literature on MC focuses on its impact on HCE. Based on the Rand Health Insurance Experiment, Manning et al. (1984) studied the effect of MC on the utilization of health care services and on the level of HCE. They had randomly assigned a group of 1,580 persons to receive care free of charge from either a fee-for-service physician of their choice (representing conventional care) or a physician participating in a Health Maintenance Organization (HMO, representing MC). In addition, a group of 733 individuals, already enrolled in a HMO, constituted a control group. The crucial innovation of this study was that participants were assigned to plans, which served to avoid risk selection effects, causing healthier individuals to enroll in MC plans. Both groups enrolled in the MC plan had 40 percent fewer inpatient admission levels than those assigned to the conventional insurance plan. Their total HCE was about 25 percent lower than under conventional care.

Cutler et al. (2000) analyzed the effect of MC on the cost of treatment of one particular disease. They compared the treatment of heart disease in Health Maintenance Organizations (HMOs) and traditional insurance plans using two datasets from Massachusetts. For the HMOs they found 30 to 40 percent lower HCE than for traditional plans, mainly due to differences in unit prices. They concluded that MC may yield substantial increases in measured productivity relative to traditional insurance. Using Swiss panel data, Lehmann and Zweifel (2004) were able to distinguish cost savings due to risk selection and due to innovation effects. They found some evidence of risk selection effects, which however, accounted for only one-third of the cost advantage in the case of HMOs, with the remainder attributed to innovation effects.

This paper follows a more comprehensive (but more descriptive) approach by assessing the contribution of MC to the performance of an entire health care system. Performance is measured using five standard criteria developed for the assessment of an economy. They are (1) matching of consumer preferences, (2) technical efficiency, (3) adaptive capacity, (4) dynamic efficiency, and (5) a rent-free distribution of income that provides incentives for producers to attain criteria (1) through (4). These criteria are applied to the three contractual relationships typically characterizing a health care system, viz. (a) between the insured and patients and health insurers (the government as it were in the case of National Health Service-type systems); (b) between insurers (the government, respectively) and health care providers; and (c) between the insured and patients and health care providers. The countries to be analyzed according to these five criteria and three contractual relationships are Germany, the Netherlands, and Switzerland. This choice can be justified for the following reasons. First, all three are insurance-based, which facilitates the comparison. Second, elements of MC were introduced in all three countries during the last few years. Third, the Netherlands underwent an important reform of their health care system in 2006, which allows to test the hypothesis that the contribution of MC to system performance depends on the institutional framework.

The paper is organized as follows. Section 4.2 contains the definition of MC and an explanation of the criteria for measuring performance. In Section 4.3, these criteria are applied to the three contractual relationships of a health care system before the introduction of MC. Finally, the contribution of MC to the performance of the three health care systems is assessed in Section 4.4 by applying the criteria to the

contractual relationships after the introduction of MC. The last section concludes.

4.2 Analytical framework

4.2.1 The common building block of health care systems and the scope of Managed Care

The common building block of all health care systems is the relationship between the patient (the principal in the economic theory of contract) and the physician (the agent). Patients experience a significant informational disadvantage, causing them to delegate decision-making authority to the physician. In particular, they may at best observe the outcome of a treatment, but not physician effort. For the physician, however, additional effort is costly, at the very least in terms of leisure forgone. This fact alone prevents physicians from being a perfect agent of their patients. Generally, physicians will set their effort at a level they consider optimal from their own point of view. Since it is in general impossible for the patient to find the payment function inducing the optimal treatment effort by the physician, there is scope for complementary agents who promise to mend the physician-patient relationship (Zweifel, 1998).

In Germany, the Netherlands, and Switzerland health insurers (in Germany medical associations as well) represent the dominant complementary agents. However, complementary agents induce new information asymmetries leading to moral hazard and adverse selection effects. Specifically, insured patients tend to consume more health care services than medically necessary. The objective of MC is therefore to rearrange the relationship between these three players in order to mitigate information asymmetries and enhance efficiency as well as to optimize the allocation of the health care resources used (Finsterwald, 2004).

Different forms of MC exist, including Health Maintenance Organizations (HMOs), Preferred Provider Organizations (PPOs), and Independent Practice Associations (IPAs). They integrate insurance and provision of health care services, however to a different extent. This integration is designed to reduce moral hazard effects between health care service provider and insurer, to optimize the use of health care resources (e.g. by avoiding double visits), and to better structure treatment processes. To this

end, MC organizations listed above apply different instruments, such as selective contracting, gatekeeping, and disease management (Amelung, 2007, Felder, 2003). In the following we will analyze the impact of MC on the performance of health care systems. The criteria used for evaluation are presented below.

4.2.2 Evaluation criteria

The five criteria listed below were originally developed for the assessment of an economy in general. Zweifel (2006) adapted them to the health care sector in the following way.

1. Matching of consumer preferences: Health care services should match the preferences of the insured, who are assumed to decide about the types of medical care that should be covered by insurance before they are ill.

2. Technical efficiency: The health care services that are provided according to criterion (1) should be produced at least cost.

3. Adaptive capacity: The insured as well as insurers and health care providers should adapt e.g. to population ageing or to medical technical change.

4. Dynamic efficiency: The health care sector should have an optimal mix of product innovation (goods with changed characteristics that may fetch a higher price) and process innovation (unchanged goods at lower cost and price). In general, insurance-induced moral hazard biases this mix in favor of product innovation (Zweifel et al., 2009).

5. Income distribution according to performance: Providers of health care services should not be able to enjoy monopolistic rents (e.g. incomes that exceed the amount that is necessary to keep them in their current activity). Rents jeopardize the attainment of the other four criteria because providers lack the incentive to make the pertinent efforts.

These criteria are applied to the three contractual relationships characterizing a health care system defined in section 2.1. Each time, the issue is whether MC contributes to the improvement of the contractual relationship in the light of the five criteria.

4.3 The contractual relationships prior to Managed Care in Germany, the Netherlands, and Switzerland

First, the health care systems of Germany, the Netherlands, and Switzerland will be described prior to MC using criteria (1) to (5). This is somewhat difficult for the Netherlands since some MC elements have been part of the health care system for a long time.

4.3.1 The contractual relationship between consumers and health insurers

In all the three countries, consumers can choose among different health insurers. However, insurance policies are highly regulated. The government not only limits the range of admissible premiums (contribution rates, respectively in Germany), but also the treatments to be covered. Only in the Netherlands, cost-effectiveness constitutes a criterion for the inclusion or exclusion of benefits (Schreyoegg et al., 2005). Dutch and Swiss health insurers have more freedom to launch different types of insurance policies than their German counterparts (Schut and de Ven, 2005 Becker et al., 2007). In Germany, only private health insurers have the right to differentiate their products (Jacobs and Schulze, 2006).

Clearly, the obligation for health insurers to offer largely uniform insurance policies makes it difficult to match consumer preferences, to quickly adapt to changes of the economic environment, and to sustain technical and dynamic efficiency. Therefore, criteria (1) through (4) are violated (see Table 4.1). However, competition for consumers has been enough in the three countries to prevent the creation of rents [criterion (5)] (Zweifel, 2006).

4.3.2 The contractual relationship between health insurers and health care providers

In Germany and Switzerland, health insurers are subject to an 'any-willing-provider' clause, i.e. they are forced by law to contract with every approved physician. In the Netherlands, selective contracting has been possible since 1994; however, health

Table 4.1: Main violations of performance criteria prior to MC

	Consumers-Insurers	Insurers-Providers	Consumers-Providers
Germany	criteria (1) through (4)	criteria (1), (2), and (5)	criterion (2)
Netherlands	criteria (1) through (4)	criteria (1) and (5)	criterion (1)
Switzerland	criteria (1) through (4)	criteria (1), (2), and (5)	criterion (2)

insurers have been making very limited use of this right up to the present (Baur et al. (2001)). Health insurers also lack the right of negotiating differentiated, incentive-compatible modes of physician remuneration in the countries analyzed. In Germany, the association of social health insurers and contract physicians negotiate both the global medical budget and the nationwide fee schedule[1] (Busse, 2004). In the Netherlands, insurers have the choice between paying (primary care) physicians either fee-for-service or using capitation (i.e. a fixed amount per enlisted patient). Swiss health insurers must apply Tarmed (Tarif médical), a nationwide fee schedule. For hospital services, German health insurers are subject to a nationwide fee schedule (Pflegesatzverordnung). Dutch insurers have some negotiating leeway, which is however constrained in several ways. In Switzerland, they are confronted with cantonal hospital associations. For pharmaceuticals, all three countries impose a national benefit list along with regulated prices.

Obviously, collective contracting and uniform payment schedules in Switzerland and Germany violate criteria (1), (2), and (5). However, criterion (2) is satisfied to a higher degree in the Netherlands due to more flexibility with regard to modes of payments.

4.3.3 The contractual relationship between consumers and health care providers

In Germany and Switzerland, patients can choose their preferred physicians without any limitation. In the Netherlands, they are obliged to see a primary care physician first, which may not be in accordance with consumer preferences [criterion (1)].

Table 4.1 summarizes the rough overall assessment of the German, Dutch, and Swiss health care system prior to MC.

[1]Medical associations and health insurer associations distribute the budgeted amount proportionally accord-ing to billed activity between the primary care physicians of a given Land.

4.4 Assessing the contribution of Managed Care to the performance of the German, Dutch, and Swiss health care systems

This section is devoted to an assessment of the contribution MC makes to the performance of the German, Dutch, and Swiss health care systems. The scales used will be 2 points if MC fully contributes to the attainment of the criterion, 1 point if MC partially contributes, and 0 points if it does not contribute to the attainment of the criterion. Points will be simply added to obtain a total score. For each country, the assessment focuses on the MC element that is most prominent, e.g. Disease Management Programs in the case of Germany.

4.4.1 The contribution of Disease Management Programs to the performance of the German health care system

The Laws on Health Insurance of 2000 and 2004 paved the way for MC in Germany (The Federal German Ministry of Justice, 2006). The governments' objective is to foster the integration of hitherto strictly separated ambulatory and hospital care[2]. In the MC setting, sickness funds[3] are allowed to selectively contract with physicians without the involvement of medical associations. Moreover, alternative forms of payment, including capitation can be implemented. The government promotes three different types of MC in particular, viz. Medical Care Centers (Medizinische Versorgungszentren), Independent Practice Associations (Hausarztmodelle), and Disease Management Programs (Strukturierte Behandlungsprogramme) (Greiner, 2005, Busse, 2004).

The effects of MC will be illustrated for the Disease Management Programs (DMPs). They have been developed to improve quality and cost-effectiveness of treatment received by the chronically ill. So far the government has defined DMPs for four chronic diseases, diabetes, breast cancer, asthma, and coronary heart disease. The sickness funds receive payments out of the risk adjustment scheme[4] for every individual enrolled in a DMP (The Federal German Ministry of Health, 2007, Wiechmann, 2003).

[2]Therefore, MC is known as integrated care (integrierte Versorgung) in Germany.

[3]In Germany statutory sickness funds act as social health insurers.

[4]The risk adjustment scheme is currently based on age, sex, gender, and the four DMPs officially implemented.

It was hoped that connecting DMPs with the risk adjustment scheme would provide a stimulus for sickness funds to attract chronically ill people rather than eschewing them as high risks.

Relationship between consumers and insurers

Preferences of patients are not considered in the definition of DMPs. However, participation in DMPs is not mandatory but offers an additional choice. Therefore matching of consumer preferences is slightly improved. There is little reason to expect that the chronically ill will obtain their treatment at lower cost because the DMPs do not provide incentives to health insurers or providers for a better coordination of care. Therefore, static efficiency is not enhanced. The Government determines the design of DMPs, e.g. it decides (using lengthy procedures) which chronic diseases are included. This does not improve adaptive capacity of the system. However, sickness funds have incentives to support cost-reducing process innovation because the payment of the risk adjustment scheme is fixed, putting them at risk for exclusive cost of treatment. This serves to redress the balance between product and process innovation somewhat. These considerations may justify the entries in the second column of Table

Relationship between insurers and providers

Thanks to selective contracting, sickness funds are supposed to become prudent purchasers on behalf of their clients. One would expect them to contract only with those physicians exhibiting a favorable cost-benefit ratio in treatment of chronic illness. However, physicians participating in DMPs lose their autonomy in medical decision-making. Beside many other regulations, they must follow treatment guidelines and document the whole treatment process electronically. On the whole, DMPs are unattractive to physicians, who continue to have the option of billing fee-for-service. Therefore the DMPs do little to increase the degree of competition between health care providers and hence to improve the matching of consumer preferences, technical efficiency, adaptive capacity, or the avoidance of monopolistic rents. At least, the DMP guidelines may induce providers to focus more on process innovation, motivating the entry +1 for criterion (4) in the second column of Table 4.2.

Table 4.2: Contribution to performance of DMPs in Germany

Contractual relationships / Criteria	Consumers Insurers (1)	Insurers Providers (2)	Consumers (3) Providers (3)	Σ (max= 6 per item) (4)
(1) Matching of consumer preferences	1	0	0	1
(2) Technical efficiency	0	0	0	0
(3) Adaptive capacity	0	0	0	0
(4) Dynamic efficiency	1	1	0	2
(5) Income distribution according performance	0	0	0	0
Σ(max= 10 per item)	2	1	0	3

0= No change; 2= improvement; 1= partial improvement. Shaded fields= MC helps to alleviate a shortcoming noted in Table 4.1.

Relationship between consumers and physicians

In general, German consumers have a free choice of physicians but are expected to visit the hospital recommended by their primary care physician. However, usually primary care physicians place no restrictions on the hospital choice (BMJ, 2006). In contrast, DMP-patients are constrained to select a participating physician. This limitation is hardly compensated; in particular, there is almost no reduction in the rate of contribution. But at least consumers are not forced to participate, justifying the zero entry for criterion (1) in the third column of Table 4.2. Likewise, attainment of criteria (2) through (5) remains unchanged.

In sum, the contribution of DMPs to the performance of the German health care system remains limited (3 out of 30 points, see the last column of Table 4.2). The main violations of the criteria prior to the introduction of MC (see shaded fields in Table 4.2) could not be offset, except for the criteria (1) and (4) relating to the relationship between consumers and insurers. This limited contribution is mainly due to comprehensive and uniform regulation stifling any innovative action by sickness funds. For the same reason, selective contracting does not produce the expected benefits for consumers. Thus, DMPs fail their main objective, viz. improved coordination and quality of treatment provided to the chronically ill. 4.2.

4.4.2 The contribution of the Independent Practice Association to the performance of the Dutch health care system

The Dutch government implemented radical market reforms with the Health Care Act of 2006, the main objective being to increase efficiency by promoting competition in all three contractual relationships. The act created a level playing field between private and social health insurers, who converted to for-profit status. By March 2006, every Dutch citizen had to explicitly choose an insurer and a policy. Competitive pressure is expected to make better use of already existing MC tools, viz. selective contracting and gatekeeping. In contrast to their German and Swiss counterparts, Dutch health insurers can selectively contract with physicians (and, to a far more limited extent, hospitals). Gatekeeping, requiring patients to first contract their primary care physician, was also established practice prior to the 2006 reform.

The effects of MC will be discussed focusing on the Independent Practice Association (IPA). The IPA is a network of primary care physicians who agree to act as gatekeepers. The main objective is to use medical care efficiently, e.g. by preventing unnecessary hospitalisations (Douven and Pomp, 2007, den Exter et al. (2004)).

Relationship between consumers and insurers

An IPA might endanger the matching of consumer preferences since patients have to visit a primary care physician first. However, health insurers are more likely to become prudent purchasers on behalf of their costumers, using their freedom to contract with physicians who match the preferences of their members. Competition for customers also forces health insurers to pass on savings to their clients and to adapt quickly to changes of the economic environment. These considerations justify the positive entries for criteria (1) through (3) in the first column of Table 4.3. In contrast, dynamic efficiency and income distribution according to performance are not attained to a higher degree because the mix between product and process innovation and provider competition for patients remains unchanged.

Relationship between insurers and health care providers

Profit-maximizing health insurers seem to be at a first glance only interested in providers that keep down cost. However, they need to contract with providers who respect patient preferences to be successful since customers can switch insurers. Insur-

Table 4.3: Contribution to performance of IPA in the Netherlands

Contractual relationships / Criteria	Consumers Insurers (1)	Insurers Providers (2)	Consumers (3) Providers (3)	Σ (max= 6 per item) (4)
(1) Matching of consumer preferences	1	1	0	2
(2) Technical efficiency	2	2	2	6
(3) Adaptive capacity	2	1	1	4
(4) Dynamic efficiency	0	1	0	1
(5) Income distribution according performance	0	2	0	2
Σ(max= 10 per item)	5	7	3	15

0= No change; 2= improvement; 1= partial improvement. Shaded fields= MC helps to alleviate a shortcoming noted in Table 4.1.

ers' freedom to contract exposes physicians to an increased intensity of competition. Because physicians do not have the alternative of contracting outside the IPA (unlike in Germany and Switzerland), freedom to contract clearly serves to enhance technical efficiency. Moreover, beside cost competition there is still scope for quality competition, making physicians adapt quickly to changes of the economic environment [criterion (3)]. They are also more inclined to adopt cost-saving process innovation since many are paid a capitation, shifting the risk of high treatment cost on their shoulders. Finally, physicians with an unfavorable cost-benefit ratio have difficulty striking contracts with insurers. On the whole, these considerations motivate the entries of the second column in Table 4.3.

Relationship between consumers and health care providers

Dutch patients do not have direct access to specialists. The restriction to see the gatekeeper first clearly is not compatible with a matching of consumer preferences. On the other hand, it does enhance technical efficiency by a more coordinated treatment process, e.g. by avoiding double visits. Since the guidelines are voluntary, physicians can increase their chances by adopting them, increasing the adaptive capacity of the system. However, the IPA does not contribute to an increased satisfaction of criteria (4) and (5). In all, the IPA of the Netherlands achieves 15 out of 30 points and therefore contributes considerably to a higher performance of the Dutch health care system. Especially the criteria violated prior to the introduction of MC are now satisfied to a higher degree (see shaded fields in Table 4.3). These are criteria (1) through (3) in the relationship between

consumers and insurers and criteria (1) and (5) in the relationship between insurers and providers. The highest score comes from the relationship between insurers and health care providers (7 out of 10 points). Among the five criteria, technical efficiency benefits most, mainly due to selective contracting (6 out of 6 points). Apparently, the changed institutional setting of the Dutch health care system causes the contribution of MC to performance to be higher than the institutional setting of Germany.

4.4.3 The contribution of Health Maintenance Organizations to the performance of the Swiss health care system

The new Law on social health insurance (KVG), effective 1996, established MC options, which had been introduced to the Swiss health care system since 1993. It enables health insurers to selectively contract with physicians (but not hospitals for the mandatory basic package). The MC alternatives offered include physician networks (similar to the afore-mentioned IPAs), Preferred Provider Organizations (PPOs), and Health Maintenance Organizations (HMOs). In the following, focus will be on the HMOs when discussing the effects of MC on the performance of the Swiss health care system. The typical Swiss HMO takes the form of capitated local group practices, with physicians as salaried employees (Lehmann and Zweifel, 2004).

Relationship between consumers and health insurers

In Switzerland, HMOs constitute an alternative to conventional fee-for-service policies. They allow consumers to voluntary limit their choice of physician in exchange for a lower insurance premium. However, government regulation prevents the premium reduction from exceeding 20 percent for the first five years of contract life. This is not sufficient for the average Swiss consumer since market experiments show that restrictions of the freedom of physician choice have to be compensated with one-third of average premium (Zweifel et al., 2006). The expression of individual preferences therefore is not quite perfect (see entry in the first column of Table 4.4). Swiss HMOs do provide health care services at up to 63 percent lower cost than conventional fee-for-service. About one-third of the amount is due to risk selection effects, while two-thirds can be attributed to changed incentives (Lehmann and Zweifel, 2004). This prevents HMOs making a more substantive contribution to technical efficiency [see the entry for criterion (2) in Table 4.4]. MC was also expected to foster product innovation in health insurance. However, prior to the introduction of MC, Swiss policy makers put a risk

adjustment scheme in place. Such a scheme binds insurers with an above-average share of young enrollees pay into the scheme. While this attempts to alleviate the problem of risk selection it in fact punishes innovators. In this way, insurers adaptive capacity fails to be enhanced. Next, MC seems to have encouraged switching by consumers to a rate of around 10 percent by 2006 (Federal Council of Switzerland, 2007). In their fight for market share, insurers can be expected to improve the cost-benefit ratio of their products, thus contributing to dynamic efficiency. However, with uniform premiums also imposed on MC alternatives, many insured are free to consume medical services without any financial consequences, which is not compatible with a no-rent (net) income distribution. These considerations lead to the remaining entries in the first column of Table 4.4.

Relationship between health insurers and health care providers

Insurers' obligation to contract with any willing provider is not compatible with a matching of consumer preferences, technical efficiency, and an income distribution devoid of rents. However, this "any-willing-provider" clause is not applicable to MC options that permit insurers to select physicians with a favorable cost-benefit ratio. With only a small share of the population enrolled in MC alternatives (18 percent in 2006, one-third of which in HMOs; see Federal Council of Switzerland, 2007), physicians still can easily revert to conventional fee-for-service with its "any-willing-provider" clause. Therefore, MC only partially enhances attainment of criteria (1), (2), and (4) while not affecting adaptive capacity [criterion (3)] and rent-free income distribution [criterion (5)].

Relationship between consumers and health care providers

Consumers signing up for a HMO accept limited physician choice. However, they do this voluntarily. As long as MC is not the dominant type of health care provision, criterion (1) is not violated. The remaining criteria (2) through (5) are not affected, mainly because patients were able to choose their physicians freely prior to MC. Their choice of hospital is restricted to the canton of residence (except if covered by supplementary insurance), and MC has not changed this. In all, Swiss HMOs contribute to a higher performance of the Swiss health care system (6 out of 30 points, see Table 4.4). As for the Netherlands, main improvements are found especially for previously violated criteria (see shaded fields Table 4.4). These are criteria (1), (2), and (4) for the relationship between consumers and insurers and criteria (1), (2), and (5) for the

Table 4.4: Contribution to performance of HMOs in Switzerland

Contractual relationships / Criteria	Consumers Insurers (1)	Insurers Providers (2)	Consumers (3) Providers (3)	Σ (max= 6 per item) (4)
(1) Matching of consumer preferences	1	1	0	2
(2) Technical efficiency	1	1	0	2
(3) Adaptive capacity	0	0	0	0
(4) Dynamic efficiency	1	0	0	1
(5) Income distribution according performance	0	1	0	1
Σ(max= 10 per item)	3	3	0	6

0= No change; 2= improvement; 1= partial improvement. Shaded fields= MC helps to alleviate a shortcoming noted in Table 4.1.

relationship between insurers and physicians. Matching of consumer preferences and technical efficiency [criteria (1) and (2)] benefit most of the introduction of MC. However, government regulation such as the "any-willing-provider" clause in conventional medicine and for hospitals, prevent MC from making a more substantial contribution to the Swiss health care system.

4.5 Conclusions

This article proposes an innovative approach to analyze the impact of MC on health care systems. Rather than just analyzing the effects of MC on HCE, it considers economic criteria, viz. matching of consumer preferences, adaptive capacity, dynamic efficiency, and income distribution according to performance, which determine the performance of a health care system.

The contribution of MC to the performance of the German health care system remains limited (3 out of 30 points). The DMPs, designed to improve the quality and cost-effectiveness of treatments for chronically ill people, cannot fulfill their expectations. Government regulation such as the uniform design of the DMPs, the bureaucratic requirements for physicians, and the loss of free physician choice without adequate compensation, prevent DMPs from making a substantial contribution to the German health care system.

In the Netherlands, MC contributes considerably to a higher performance of the Dutch health care system (15 out of 30 points). MC together with the Health Care Act of 2006 liberalized the relationship between health insurers and health care service providers, making health insurers prudent purchasers of health care services on behalf of their clients. Therefore the main improvements are found for the relationship between health insurers and health care service providers (7 out of 10 points) and for the second criterion, technical efficiency (6 out of 6 points). In Switzerland, MC contributes to a higher performance of the health care system, but to a lower degree than in the Netherlands (6 out of 30 points). Improvements are found for the relationships between consumers and health insurers, and health insures and health care service providers (both 3 out of 10 points). Among the five criteria, matching of consumer preferences [criterion (1)] and technical efficiency [criterion (2)] receive the highest score (2 out of 6). However, regulations such as the risk adjustment scheme, the uniform benefit package, or the 'any-willing-provider' clause governing the relationship between health insurer and health care provider, prevent MC from making a more substantial contribution to the Swiss health care system.

Finally, the findings suggest that MC depends on the institutional setting. The more freedom to contract between consumers, health insurers, and health care service providers, the greater the contribution of MC to the health care system. However, further research is necessary to test this hypothesis.

Acknowlegements

The authors would like to thank Karolin Leuckert, Boris Krey, Maria Trottmann, Philippe Widmer, and participants of the 2008 Annual Meeting of the SSES for helpful comments.

Chapter 5

Fine-Tuning of Health Insurance Regulation: Unhealthy Consequences for an Individual Insurer

Johannes Schoder, Michèle Sennhauser and Peter Zweifel

Published in

International Journal of the Economics of Business, 17(3), 2010, 313-327.

5 Fine-Tuning of Health Insurance Regulation: Unhealthy Consequences for an Individual Insurer

5.1 Introduction

When premiums are mandated to be independent of risk, competitive health insurers have an incentive to select clients whose future expected health care expenditure (HCE) does not exceed their contribution. This consideration has induced secondary regulation in the guise of risk adjustment (RA) schemes. Basically, RA makes insurers with an above-average share of favorable risks pay into a fund, whose proceeds are used to cross-subsidize those insurers with many unfavorable risks. The design of an optimal RA formula is a widely discussed topic (see for example Lamers, 1999, Ellis and Van de Ven, 2000, Glazer and McGuire, 2002, Lamers and Van Vliet, 2003a, Lamers and Van Vliet, 2003b, Van de Ven et al., 2004, Beck et al., 2006, Jack, 2006, Zweifel and Breuer, 2006, and Van de Ven and Schut, 2007). The RA formulas for Medicare in the United States and the Netherlands are being refined continuously (see e.g Douven, 2007 and Calfo, 2009). However, so far the consequences of this fine-tuning of regulation for the risk management (RM) of insurers seem to have been neglected.

This contribution contains a case study from Switzerland, a country that relies on competitive health insurance in a way similar to the US and the Netherlands. A RA scheme was introduced in 1996, using the two criteria age and gender only. Effective 2012, the RA formula will include a third indicator of high risk, viz. "Hospitalization of more than three days or living in a nursing home during the previous year" (see Spycher, 2000). While this choice is largely dictated by service providers' refusal to pass on diagnostic information to health insurers, it does have several recommendable features in that it (1) has significant predictive power (see Holly et al., 2003 and Beck, 2004), (2) relates to a previous period so does not undermine insurers' effort

at controlling health care cost, and (3) can be measured at little administrative expense.

Refinement of the RA formula has gone much further in other countries. In the United States, the CMS hierarchical condition categories model (CMS-HCC) has been in use with Medicare since 2004. It uses diagnoses from all clinical encounters, regardless of whether they are inpatient or outpatient (see Pope et al., 2004). In the Netherlands diagnostic cost groups (DCGs) and pharmacy-based cost groups (PCGs) are used as high-risk indicators.[1] These reforms have their costs and benefits. On the benefit side, risk-selection efforts by health insurers are reduced if the net cost of medical care falling on them is increasingly equalized across risk types. Moreover, this net cost does not depend anymore on whether the insured were hospitalized or not. On the cost side, these refinements of RA not only require more accounting effort on the part of both insurers and providers but also increase proneness to error[2]. Moreover, they create incentives for up-coding diagnoses (for an explicit analysis of advantages and disadvantages in the case of United States Medicare, see Pope et al., 2000 and Kominski, 2007).

The purpose of this paper is to point out another cost of RA refinement. Indeed, it may boost payments into the RA scheme to an extent as to jeopardize the economic survival of an otherwise viable health insurer, posing a great challenge to its RM. Now insolvency and hence market exit of an insurer who only survived thanks to cream skimming may be considered to be efficiency enhancing. However, this case study deals with an innovative health insurer, who had successfully implemented Managed Care (MC) to lower rates of hospitalization. Bankruptcy of such an insurer would have to be considered inefficient.

The evidence comes from simulating payments for a particular health insurer A into the RA scheme applying the old and the new formula. These simulations predict that A's payments would have increased significantly, attaining between 9 and 13 percent of premium income. Extra payments of these magnitudes would have seriously endangered insurer A's economic survival, leading to a cumulative loss

[1] They are derived from the diagnoses related to prior hospitalization and prior use of prescription drugs, see e.g. Van de Ven and Schut (2008).

[2] In the Netherlands, the complexity of processing the data and money flows led to errors in the calculation of the ex-ante risk-adjusted capitation payments, resulting in a loss of Euro 247 million (mn.), falling on taxpayers (see Douven, 2007).

in excess of CHF 250 mn. (CHF 1 ≈ USD 1 at 2010 exchange rates) over three years. While A's RM response cannot be predicted, there are two main alternatives. One is to enlist unfavorable risks, as intended by the regulator. The other is to extend hospital stays from three to four days. This strategy would have decreased this insurer's RA payments by an estimated 11 percent in 2007. The consequences would be unhealthy for taxpayers (who subsidize hospital cost), employers (who lose workdays), and patients (who lose quality of life). While not directly transferable to other countries with competitive health insurance (such as the United States, but also Germany, Israel, and the Netherlands), the findings of this contribution convey a clear message. Seemingly minor fine-tuning of health insurance regulation has the potential of challenging an insurer's RM, with undesirable consequences for the society.

The remainder of this chapter is structured as follows. Section 5.2 describes the method for calculating risk adjustment values in general and the data basis. In the first part of Section 5.3, RA values are simulated according to the new formula and applied to insurer A. The second part of Section 5.3 analyzes the impact of this regulatory change on insurer A's RM. The chapter concludes with lessons learned from this case study and its implications.

5.2 Simulation of risk adjustment values and data basis

5.2.1 Methodology

Traditionally, analysis of RM focuses on payments between health insurers. However, this neglects the fact that payments into the RA scheme are ultimately borne by low-risk consumers while payments from the scheme benefit high-risk consumers. Economic theory has always distinguished between payers and bearers of a cost or levy, in particular in the context of an indirect tax. To see the analogy, consider current Swiss RA with two criteria age and gender only. Define $\bar{P}$ as the community-rated premium, $\bar{L}_{a,g}$, as the average HCE in one of the age-gender cells (a, g) of RA (neglecting administrative expense for simplicity), and $RA_{a,g}$ as the payment to or from the RA scheme. The premium paid by a specific individual i who is a low risk compared to the cohort

in the age-gender cell (a, g), and whose expected cost $\mathrm{E}(L_i)$ is thus below average for the specific cell can then be expressed as

$$\begin{aligned} \bar{P} &= \bar{L}_{a,g} + RA_{a,g}, \qquad \text{with } RA_{a,g} > 0 && (5.1) \\ &= \mathrm{E}(L_i) + (\bar{L}_{a,g} - \mathrm{E}(L_i)) + (\bar{P} - \bar{L}_{a,g}). && (5.2) \end{aligned}$$

This particular low risk bears, on top of his or her actuarially fair premium $\mathrm{E}(L_i)$, a cross-subsidy in favor of high risks consisting of two components. The first component is the difference between average HCE of group (a, g) and the individual's expected HCE denoted by $\mathrm{E}(L_i)$; the second, the contribution to the RA scheme $(\bar{P} - \bar{L}_{a,g})$, to be paid by the insurer. The sum of the two will be referred to as cross-subsidization values. As to the second component, the current Swiss RA formula comprises 15 age classes, starting from age 19 to 25 and continuing in 5-year steps. Thus, there are overall 30 RA categories. Since by law risk adjustment must not lead to a cross-subsidization between the 26 cantons (i.e. member states of Switzerland), the RA values are calculated yearly for each canton by the Joint Organization KVG based on data of all Swiss health insurers (see Joint Organization KVG, 2007). Adopting the insurer's point of view rather than the consumer's now, the RA values are equal to

$$RA_{a,g} = \bar{L}_{a,g} - \bar{L} \tag{5.3}$$

with $\bar{L}$ ($= \bar{P}$ in Eq. (5.1) since administrative expense is neglected) denoting average HCE in the canton's population as a whole (see Beck et al., 2006, ch. 4). Including the criterion "hospitalization"[3] changes Eq. (5.3) to

$$RA_{a,g,h} = \bar{L}_{a,g,h} - \bar{L}. \tag{5.4}$$

The subscript h is equal to 1 if a hospital stay in the previous year exceeds three days and 0 otherwise. Average HCE of the respective RA cell, $\bar{L}_{a,g,h}$, now has to be calculated for 60 instead of 30 groups, while $\bar{L}$ remains the same.

The insurer has to contribute to the RA fund for favorable risks ($\bar{L}_{a,g,h} < \bar{L}$). The RA fund uses the proceeds to cover the deficits generated by unfavorable risks ($\bar{L}_{a,g,h} > \bar{L}$).

[3]This is shorthand for "Hospitalization or living in a nursing home during the previous year of four days and more".

An insurer's total payment (V) into/from the RA fund depends on the composition of its insured over all 26 cantons (c),

$$V = \sum_{c=1}^{26} \sum_{h=0}^{1} \sum_{g=0}^{1} \sum_{a=1}^{15} RA_{a,g,h,c} \cdot n_{a,g,h,c}. \tag{5.5}$$

An insurer receives payments if $V>0$ and contributes to risk adjustment if $V<0$.

5.2.2 Data basis

For calculating the $RA_{a,g,h,c}$ values in Eq. (5.5) for a given health insurer, the cell-specific averages $\bar{L}_{a,g,h,c}$ must be known. Since $RA_{a,g,h,c}$ is not published by the Joint Organization KVG, two different sources are used to analyze the impact of the new RA formula on an individual health insurer. The first is constructed by merging individual HCE data provided by three large health insurers in order to calculate the average $RA_{a,g,h,c}$. Ideally it should be representative of all Swiss health insurers. The second data base comes from the one individual Swiss health insurer "A". Both are limited to individuals having mandatory health insurance.

Descriptive statistics

Data of the three large Swiss health insurers (out of a total of 70 serving a population of 7.5 mn.) is available for the period 2001 to 2005. The sample is well balanced with respect to gender (49.5 percent of women), and average age of adult enrollees (47.4 years in 2005, compared to 47.8 years of the adult population). The market share covered is stable across age classes, amounting to 25 percent on average. With regard to choice of contract, there is a clear trend towards higher deductibles. The three highest deductibles (CHF 1,500, 2,000 and 2,500; CHF 1 $\approx$ USD 1 at 2010 exchange rates) increased in importance from 12 to over 22 percent from 2001 to 2005, which is compared to the official figures of 13 and 23 percent very representative (santésuisse, 2010a). There is a similar trend in favor of MC contracts, reaching a share of 11 percent in 2005 (compared to the Swiss average of less than 10 percent in 2005, see Eugster et al. (2010)).

The second data source, obtained from A, covers the period 2001 to 2007. With 51.3 percent of women, the sample is almost balanced. A is one of the medium-sized health insurers in Switzerland with a market share of almost 5 percent in 2005. With 47.7

years, average age of A's adult enrollees is slightly higher than the 47.4 years of the three insurers. The clientele of A also tends towards higher deductibles. The share of the three highest deductibles (they are CHF 1,000, CHF 1,500, and CHF 2,500) exceeds the nation-wide average of 22 percent in 2005. MC contracts account for almost 35 percent (2007), double the nationwide average of 16.9 percent (santésuisse, 2010b). This most likely explains A's comparatively low rate of hospitalization (see Figure 5.3 below).[4] On the whole, A looks like an innovative insurer that encourages MC options, in conformity with stated objectives of Swiss policy makers.

Checking simulated RA payments

First, the data provided by the three large health insurers had to be checked for representativeness using the current RA formula. The values for $RA_{a,g}$ were calculated for all 30 cells along with their standard errors according to the methodology described in Section 5.2.1 and compared with the official nationwide values. The insurers on average pay for women aged 19 to 25 more than CHF 1,700 per year (see Figure 5.1 for the canton of Zurich, the leading canton of Switzerland both in terms of GDP and population, and Table 5.1 of the Appendix for all cantons). Conversely, they receive payment for over 90 year old women to the tune of some CHF 8,600. While the fit is good in general, RA contributions by the three insurers are lower than the official figures from age 61 on.

Based on the evidence, one can conclude that the three major health insurers sampled are sufficiently representative of the Swiss population to enable a simulation of the new RA formula based on their data. This conclusion is also supported by the fact that one of the three is a net recipient of payments from the RA scheme, one breaks even, and one is a net contributor to the scheme. Also note that according to Table 5.1 of the Appendix, the standard error and hence variance of RA payments increases with age, reflecting the fact that variance of HCE increases as well. This means that for a risk-averse health insurer, risk-selection effort has a high payoff if focused on older clients. By the same token, however, an insurer like A who counts on having to pay into the RA scheme permanently faces a liability characterized by great risk as its insured population ages.

[4] In the US, MC plans have achieved most cost savings by reducing inpatient hospital use (see Miller and Luft, 1997 and Bindman et al., 2005. For MC cost savings in Switzerland see Lehmann and Zweifel, 2004).

Figure 5.1: Official RA values according to age and gender, canton of Zurich, CHF (2005)

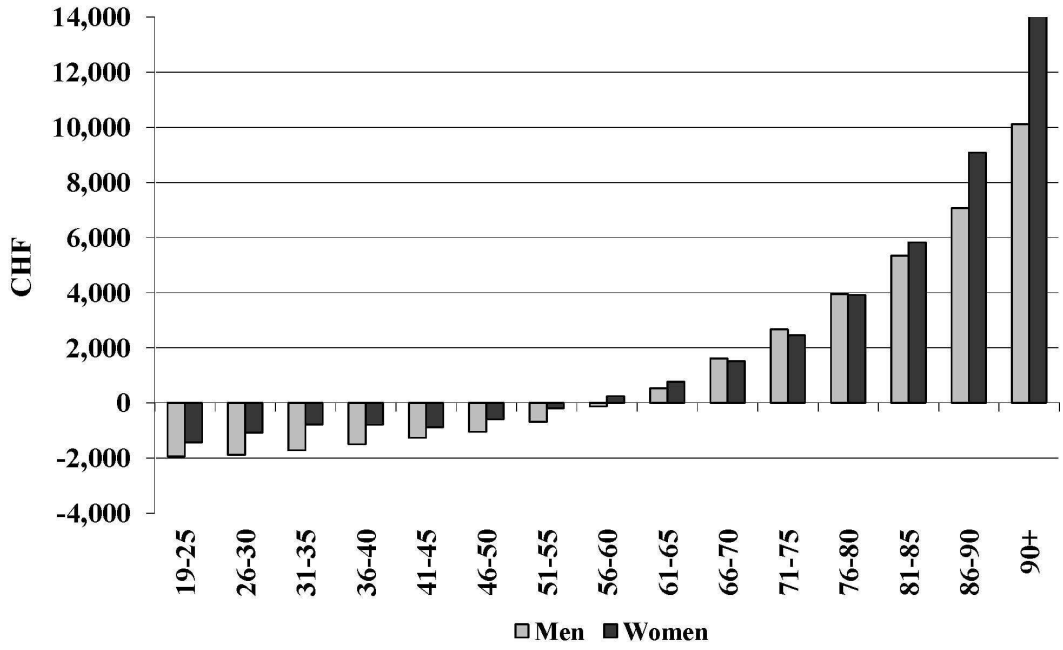
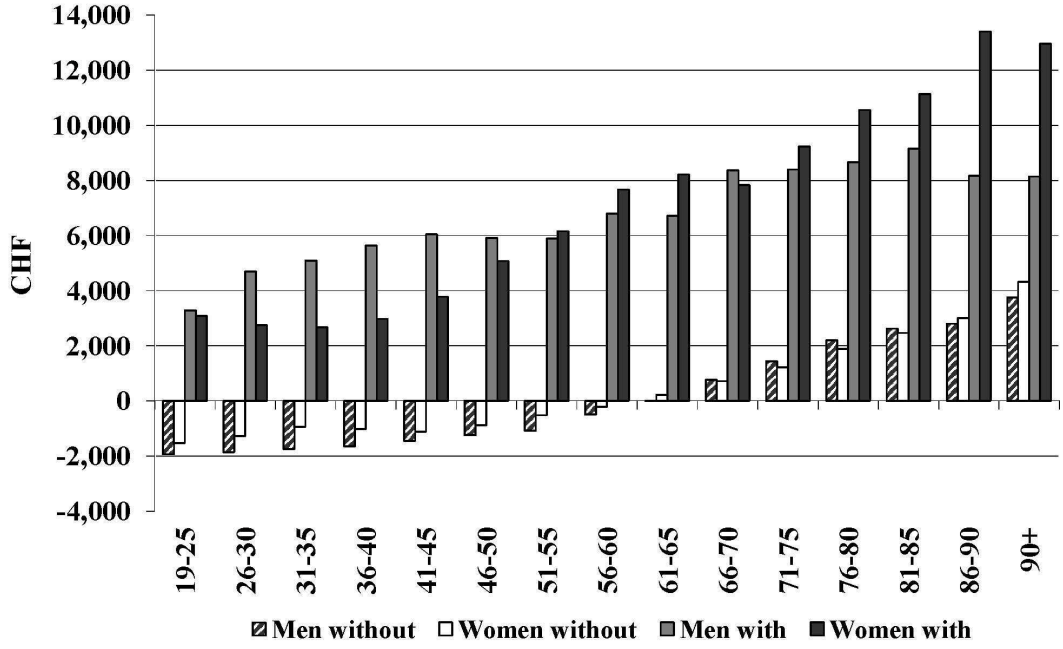


Figure 5.2: Estimated RA values with and without hospitalization according to age and gender, canton of Zurich, CHF (2005)



5.3 Simulating the impacts of the new RA formula

In this section, estimated RA values with the new RA formula including hospitalization during the previous year are presented first. Then, the impacts of the regulatory fine-tuning on health insurer A in terms of financial burden and choice of strategy are shown.

5.3.1 Risk adjustment with the new criterion

Official RA values grouped according to the additional criterion, "Hospitalization during the previous year" are not available.[5] They have been simulated using the individual HCE data provided by the three major health insurers (see Section 5.2.2). Figure 5.2 illustrates estimated $RA_{a,g,h,c}$ values for the canton of Zurich.
Comparing Figures 5.1 and 5.2 the new formula is seen to induce radical changes. First of all, it causes the amount of cross-subsidization between those without a hospital stay in the previous year to shrink considerably beyond age 70. Conversely, it causes persons with a hospital stay to be cross-subsidized regardless of age or gender. Second, and related to this, the usual age profile ceases to exist. For instance, hospitalized women in the 19 to 25 age group benefit more than the three next older groups, and at the high end, it is the aged 86 to 90 rather than the oldest that benefit most. Among men, the age profile becomes almost level beyond age 70. Third, the per capita amounts now are higher, pointing to a substantial increase in the volume of cross-subsidization. Eugster et al. (2010) simulate the effects of introducing the third criterion on the total volume of cross-subsidization for 2005. They find an increase of 40 percent, from CHF 4.13 billion to CHF 5.82 billion, or some 12 percent of Swiss HCE. Whether this is excessive or not is an issue that cannot be addressed in this paper. However, a change of this magnitude is likely to present a challenge to the RM of at least some health insurers. Whether this is the case of insurer A is the topic of the two subsections below.

5.3.2 Impacts on risk adjustment payments by health insurer A

The consequences of adding the new risk adjuster "hospitalization" for health insurer A can be simulated as follows. The volume of payments is calculated as the number of

[5] Official statistics do show RA values as "RA payments between consumers", but only according to the current RA formula (see Joint Organization KVG, 2007).

A's customers in a RA cell[6], times the estimated RA value pertaining to that RA cell, and adding up (see Eq. (5.5), Section 5.2.1). These calculations are performed using the old and the new RA formula for the years 2005 to 2007. They allow to "postdict" the consequences the new RA formula would have had if already in effect. The results are striking.

- Total payments of A into the RA scheme increase substantially. Under the old formula, they amount to CHF 24.2 mn. in 2005, corresponding to 3 percent of premium income. Had the new RA formula already been in effect, they would have reached CHF 101.6 mn., amounting to no less than 13 percent of premium income. Considering that A operated at a loss of CHF 8.2 mn. in 2005, the new formula would, ceteris paribus, have caused a total loss of CHF 85.6 mn. $(= 8.2 + 101.6 - 24.2)$.

- For the years 2006 and 2007, payments according to the new RA formula are estimated to be CHF 73.5 and CHF 82.3 mn., respectively, compared to the CHF 2.6 mn. and CHF 2.3 mn. under the current RA formula. In terms of premium income, the shares would have been 9 and 13 percent, respectively, resulting in losses of CHF 54.8 and CHF 86.2 mn., ceteris paribus.

- Payments of A into the RA scheme increase in all cantons. In some, A even turns from receiver into payer, such as in the cantons of Vaud (VD) and Geneva (GE). This precludes a regional restructuring of A's business as a possible RM response; for this reason, this alternative will not be discussed in Section 5.3.3 below.

Arguably, these developments would have jeopardized A's economic survival. Starting with the underwriting result, the combined ratio (defined as loss payments plus administrative expense plus RA values relative to premium income) was very close to 100 percent over the time period considered, viz. 102.3 (2005), 99.8 (2006), and 100.3 percent (2007).[7] This is not fatal as long as the insurer is making enough profits from capital investment (see e.g. Zweifel and Eisen, 2003, ch. 5), which was indeed the case in 2007. However, the new RA formula would have caused the combined ratio to attain 111.9 (2005), 107.5 (2006), and 110.7 percent (2007) respectively, amounts that could not have easily been compensated by profits from capital investment. According to Browne and Hoyt (1995), who analyze market predictors of insolvencies in US

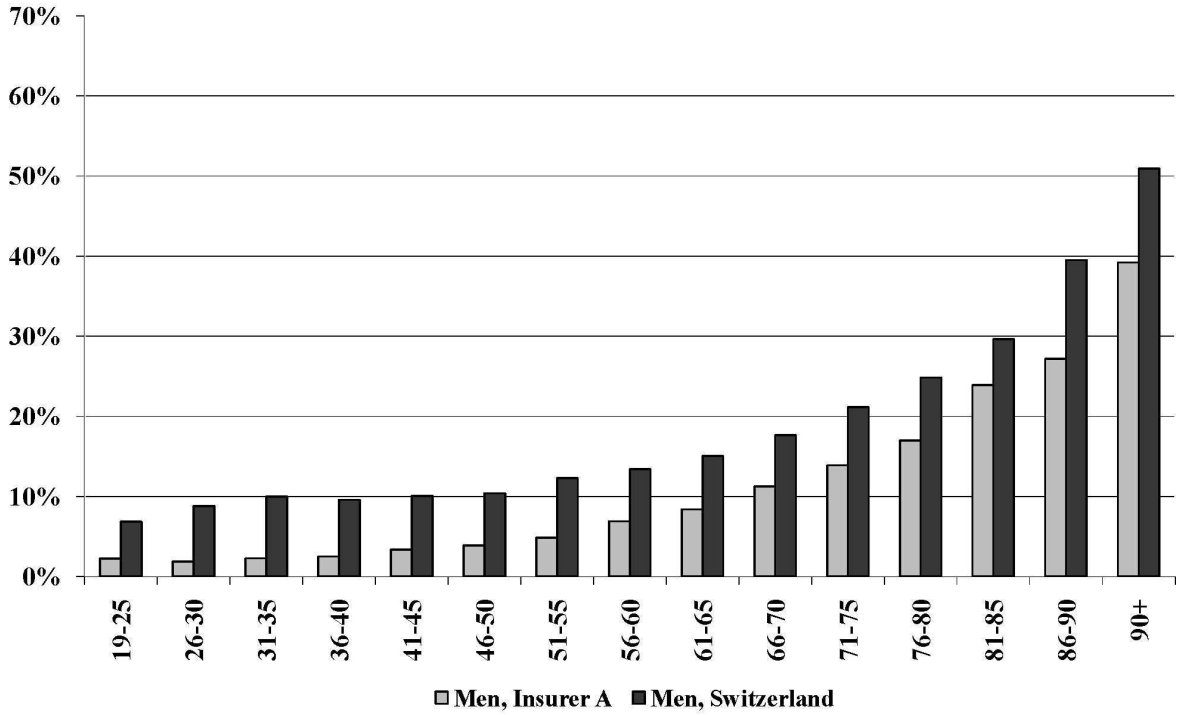
[6]For added precision, calculations are based on months of contract life.

[7]The expense ratio was 5.6 (2005), 5.9 (2006) and 5.6 percent (2007), which is average for Swiss statutory health insurers.

property-liability insurance between 1970 and 1990, a 5 point increase of the combined ratio causes the insolvency rate to increase by roughly 22 percent. Even if this result cannot be directly applied to health insurers operating in a different country, a 10 point hike in the combined ratio must substantially increase the insolvency risk of an insurer who has limited reserves. The ordinance on health insurance (Federal Council of Switzerland, 2003) requires insurers to hold reserves as a function of enrollment. With more than 150,000 insured, A currently must have reserves amounting to 10 percent of annual premiums (Santésuisse, 2009). If A would have used its reserves to make up for the predicted loss of 2005 under the new RA formula, this ratio would have fallen to around 5 percent. The predicted loss of 2006 and 2007 would have wiped out its reserves altogether.

The insolvency of an insurer could be the result of lackluster performance and hence of little importance to the economy as a whole. However, this does not seem to be true of insurer A. It did incur a loss in 2005 but was able to turn this into a surplus for the years 2006 and 2007. In addition, its high predicted payments into RA under the new RA formula are due to its low hospitalization rates (see Figure 5.3). For men

Figure 5.3: Hospitalization rate, insurer A vs. simulated nationwide values, men (2005)



(gray bars), they are significantly lower than the Swiss average (black bars) across all age groups (women similar but not shown). While successful risk selection cannot be excluded completely as an explanation, the evidence points in a different direction.

First, as stated in Section 5.2.2, the younger age classes and men are only slightly over-represented. A systematic risk selector would have significantly higher market shares in this age segment. Second, MC contracts (designed to prevent or shorten hospital stays) attain a share of 35 percent in 2007, way above the Swiss average of 16.9 percent. At the same time, insurer A's distribution of MC contracts across age classes does not systematically differ from that of the representative three insurers. Third, total HCE per enrollee and its age profile are quite similar between insurer A and the three others, speaking against across-the-board risk-selection effort on the part of the insurer A. By way of contrast, Figures 5.4 and 5.5 reveal a marked difference with regard to the cost of inpatient and outpatient care. Starting with the age group 51 to 55 but especially beyond age 81, insurer A is markedly below the simulated nationwide benchmark (Figure 5.4). Now this could still be due to risk-selection efforts cleverly targeted at the healthy elderly. In that case, however, one would also expect insurer A's cost for outpatient care to be comparatively low in the higher age groups. Yet Figure 5.5 shows that insurer A's cost of outpatient care per enrollee is higher than that of the three representative insurers, and particularly so in the high age groups.

These findings lend credibility to insurer A's claim to have implemented MC in general and home care instead of hospital care specifically for the elderly. This has positive effects not only for the individual patient whose quality of life is higher, but also for the economy as a whole. Indeed, the cost of outpatient care evidenced in Figure 5.5 is only one-half of the true value since the cantons finance roughly 50 percent of hospitals' operating cost. Implementation of MC concepts thus provides relief to taxpayers. Hence, rather than acting as a "cherry-picker", insurer A seems to be among the foremost in conforming with stated objectives of Swiss health policy, i.e. to achieve savings through MC. Insolvency of such an insurer caused by a change in the RA formula can be justifiably qualified as regulatory failure.

Figure 5.4: Outpatient cost, insurer A vs. simulated nationwide values, CHF (2005)

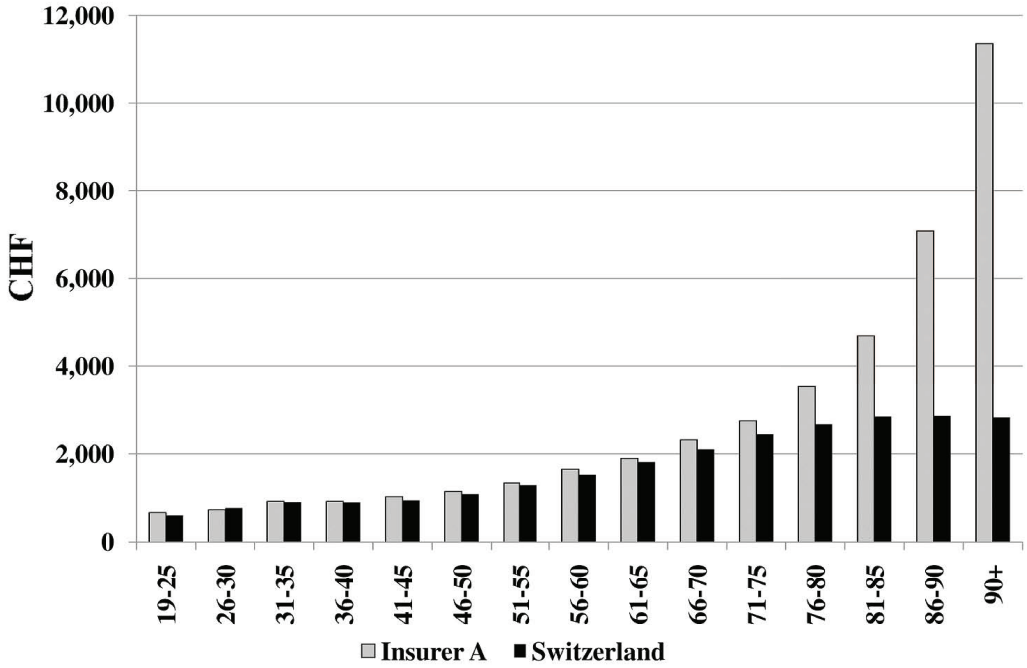
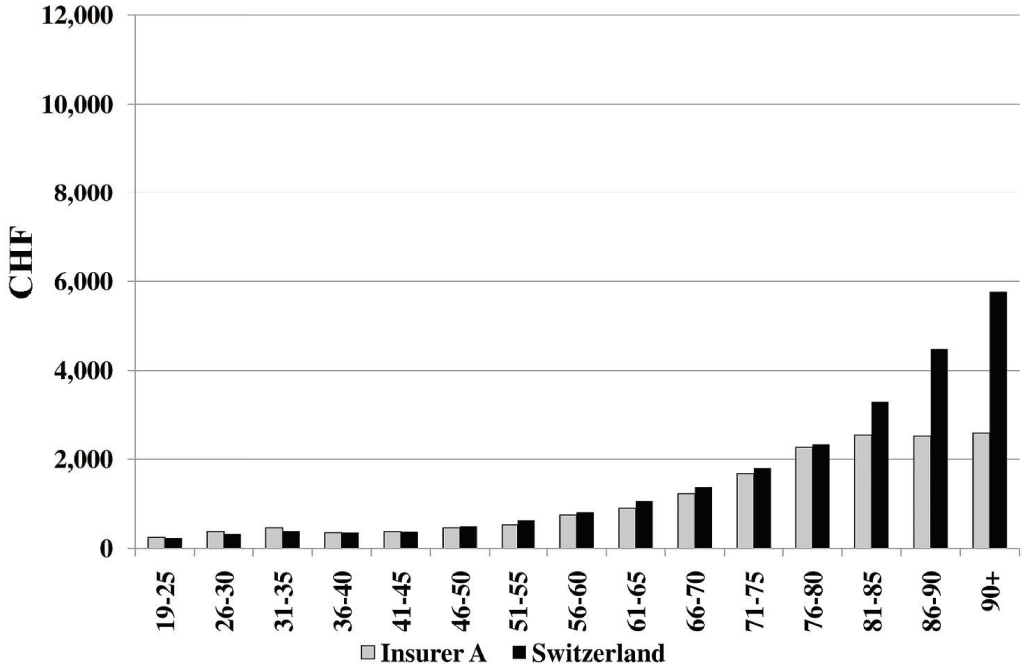


Figure 5.5: Inpatient cost, insurer A vs. simulated nationwide values, CHF (2005)



5.3.3 Impact on risk management

It is unlikely that an insurer confronted with the changes described in the preceding sections can continue with its RM strategy unchanged. The two main alternatives revolve around the two principal activities of an insurer, viz. underwriting and capital investment. Starting with the latter, the insurer could seek offsetting returns on capital investments. However, in the present state of the economy this is very difficult. In addition, capital market theory predicts that higher expected returns can only be achieved in return for more risk once the efficient frontier has been reached, a consequence that is not easily accepted by a regulator of social health insurance. The second possibility is to increase margins from underwriting either by increasing net premiums or reducing claims. Swiss statutory health insurers have to pay by law for all services included in the official list of benefits, with most prices regulated. Therefore, it is not possible to decrease insurance claims significantly. Liabilities arising from underwriting can be reduced by purchasing reinsurance; however, up to present reinsurers have not been providing coverage against RA liabilities. This leaves an increase of premiums net of RA payments as the likely RM response. Since premiums are fixed by community-rating regulation, lowering payments into the RA scheme becomes the preferred alternative.

One way to achieve this objective is to enroll more unfavorable risks, in particular persons who were hospitalized during the previous year. This is the adjustment the new RA formula was designed to bring about. The challenge to the insurer's RM now becomes to achieve more hospitalizations without incurring much additional cost. Recall that a hospitalization counts as soon as it exceeds three days. When segmenting A's HCE function according to length of stay in the hospital during the previous year, it turns out that patients with four days do not cost significantly more than those with three. Therefore, A has to weigh the once-and-for-all extra cost of a hospital day against the extra contribution from the RA scheme, which may amount to several thousand CHF (see Table 5.1 of the Appendix).

The possible reduction of RA payments can be estimated as follows. While it may not be possible to collude with the public hospitals (who obtain a per diem roughly twice the amount paid by the insurer because one-half of their extra operating cost is covered by the canton) to extend all hospital stays from three to four days, this should be possible in 50 percent of all cases. The effect of such a RM response can

be estimated with sufficient precision for the three cantons where A has the highest market share [viz. Zurich (ZH), Berne (BE), and Vaud (VD)]. There, it would have reduced RA payments by CHF 5 mn. in 2007. Extrapolating to A's entire book of business, one obtains CHF 9 mn., or 11.2 percent of the estimated CHF 82.3 mn. Savings of this magnitude would have been important enough to induce a change in RM.

The cost of this change would fall on taxpayers (who cover one half of the increased operating costs of public hospital through cantonal subsidies), employers (who bear the workdays lost), and patients (who presumably enjoy a higher quality of life outside the hospital). For this reason, reducing the length of hospital stays has been a stated goal of Swiss health policy, notably justifying the introduction of hospital payment through diagnosis-related groups by 2012 (DRGs, see SwissDRG, 2009). Thus, the fine-tuning of regulation through an improvement of the RA formula risks to burden the economy with sizable inefficiencies.

5.4 Conclusions

Regulation may pose unintended challenges to the risk management (RM) of a company. This chapter analyzes the case of health insurance, where the imposition of community-rating creates an incentive to select favorable risks. Risk adjustment (RA) schemes have been implemented in several countries such as Germany, Israel, the Netherlands, and the United States to counteract this incentive. They make insurers with an above-average share of favorable risks (indicated by age, gender, and other adjusters) to pay into the scheme, which supports insurers with an above-average share of unfavorable risks. Since its current RA formula fails to neutralize the incentive for risk selection, Switzerland will complement it in 2012 with the adjuster, "Hospitalization of more than three days or living in a nursing home during the previous year". This seemingly minor fine-tuning of regulation is shown to have a potentially fatal effect on a particular health insurer A whose payments into the RA scheme would have increased substantially between 2005 and 2007 if the new RA formula had been in effect. The reason is a low rate of hospitalization thanks to a commitment to Managed Care (MC). Therefore, A's most likely RM response would have been to increase recognized hospitalizations by increasing length of stay from three to four days, triggering extra payments from the RA scheme at a limited once-and-for-all cost of an extra hospital day. The cost of this change of RM strategy would

have been borne by taxpayers (through increased subsidies of hospitals' operating expense), employers (through lost workdays), and patients (through lower quality of life).

There are lessons to be learned for other countries who impose community-rating on competitive health insurers. First, it is practically impossible to fully neutralize insurers' risk selection incentive through an RA scheme,[8] and be it only due to their different rates of discount in estimating the present value of the benefits and costs associated with risk selection. Second, perfecting the RA formula can have unintended side effects at the level of an individual insurer that go as far as jeopardizing its economic survival in spite of innovative effort. In the case studied here, the insurer is even punished for its innovative commitment to MC. Finally, the threat of survival may well trigger adjustments in RM strategy that cause an efficiency loss to the economy as a whole.

Acknowledgments

The authors gratefully acknowledge helpful suggestions by Maria Trottmann (University of Zurich, Switzerland), Frank Lichtenberg (Columbia University, New York, USA), and an anonymous referee. Furthermore, special thanks go to the three health insurers and health insurer A that provided data for this study.

Appendix

[8]There is the perception that for all its refinement, the all-encounter RA CMS-HCC model overpays Medicare Advantage Programs (representing MCOs). We owe this interesting point to the anonymous referee.

Table 5.1: Simulated and official cross-subsidies per capita according to age and gender, CHF (2005)

Men	Average*	Std.error	Min	Max	Official value
19-25	−2,006.50	505.52	−3,006.17	−707.84	−1,963.87
26-30	−1,227.59	833.80	−2,165.91	2,287.40	−1,889.64
31-35	−900.68	678.91	−1,733.38	1,202.03	−1,771.42
36-40	−979.03	421.93	−1,749.27	247.62	−1,624.49
41-45	−828.69	351.55	−1,435.17	−40.31	−1,398.94
46-50	−543.46	465.97	−1,615.88	349.08	−1,091.94
51-55	−109.82	378.55	−977.63	714.71	−624.63
56-60	290.34	300.27	−557.57	815.53	13.40
61-65	884.74	418.34	228.53	1,648.89	771.06
66-70	1,560.60	598.50	187.69	2,464.57	1,638.40
71-75	2,535.19	548.54	982.57	3,435.54	2,873.43
76-80	3,208.98	653.35	1,884.58	4,128.30	3,845.50
81-85	4,127.79	1,361.80	1,261.52	6,983.73	4,986.30
86-90	5,286.51	1,208.24	2,752.09	7,945.75	6,880.09
90+	6,731.78	1,513.63	2,945.10	8,915.78	9,541.96
Women	Average*	Std.error	Min	Max	Official value
19-25	−1,772.99	494.20	−2,780.08	−974.44	−1,484.37
26-30	−1,024.61	461.54	−2,211.50	−311.71	−946.01
31-35	−746.06	559.49	−1,694.31	−1,125.73	−749.83
36-40	−961.00	328.45	−1,576.69	−316.11	−924.81
41-45	−965.85	279.05	−1,749.34	−535.99	−922.02
46-50	−732.01	309.04	−1,295.60	−177.44	−646.82
51-55	−442.87	268.14	−1,045.08	106.95	−235.80
56-60	−15.51	321.10	−512.16	841.85	205.36
61-65	443.65	247.14	19.55	764.95	737.31
66-70	981.80	395.53	210.13	1,603.77	1,415.39
71-75	1,982.76	446.04	758.34	2,662.32	2,385.07
76-80	3,136.84	656.22	1,838.10	4,406.12	3,671.81
81-85	4,641.23	775.55	2,788.30	6,111.25	5,596.14
86-90	6,917.12	987.66	5,115.11	8,382.98	8,486.06
90+	8,672.75	1,770.15	4,464.86	11,619.96	12,457.28

Note: * Average over all 26 Swiss cantons, 1 CHF = 0.83 USD (2007)

Chapter 6
Conclusion

6 Conclusion

The conclusion first restates the findings of each chapter and their implications followed by suggestions for possible future research.

Chapter 2 addresses an issue that has been overlooked in the production of health literature with its emphasis on flat-of-the-curve medicine. For risk-averse individuals, not only the level of health but also its variability is important. Between 1960 and 2005 uncertainty with regard to time to death decreased for the 24 OECD countries sampled. This reduction is importantly due to health care expenditure and GDP. Comparing the marginal cost in terms of health care expenditure with the willingness-to-pay values for the United States and Switzerland, we find that the benefits in terms of reduced uncertainty with regard to time to death exceed the extra cost. There are two main implications of this chapter. First, countries operating on the flat-of-the curve medicine do not necessarily waste resources, unless the variance of health status remains constant as well. Second, the reduction of uncertainty with regard to health status can also be interpreted as mortality inequality. In this respect, policy makers may also be interested in the reduction of this inequality in addition to increasing life expectancy.

Chapter 3 suggests that there are specialization gains on the regional level for the treatment of cancer. From a simple model of cancer survival we derived the testable hypothesis whether regional specialization in the treatment of cancer increases life expectancy. Specialization is measured by the number of diagnosed cancers. Using data from the National Cancer Institute's Surveillance Epidemiology End Result (SEER) we find that patients tend to survive longer in those areas where relatively more cancers of the same type exist than the US average. Possibly, a higher prevalence of cancers in some regions has led to greater accumulation of disease-specific knowledge which finally contributed to improved health outcomes. The results of our study imply that centralization of cancer treatment is justified on the ground of improved health outcomes. However, due to knowledge spillovers the level of centralization is not

necessarily limited to the hospital level (as argued in many volume-outcome studies, see e.g. Halm et al., 2002). Furthermore, since specialization is often made responsible for the growth of U.S. health care expenditure (see e.g. Baicker and Chandra, 2004) future research should relate the additional life years gained to the additional costs in order to assess whether specialization in cancer treatment is justified economically.

Chapter 4 proposes an innovative approach to analyze the impact of Managed Care on health care systems. Rather than just analyzing the effects of Managed Care on health care expenditure, it considers economic criteria, viz. matching of consumer preferences, adaptive capacity, dynamic efficiency, and income distribution according to performance, which determine the performance of a health care system. The findings suggest that Managed Care depends on the institutional setting. The more freedom to contract between consumers, health insurers, and health care service providers, the greater the contribution of Managed Care to the health care system. However, further research should be based on empirical evidence. This may come from health insurance data.

Chapter 5 analyzes the case of health insurance, where the imposition of community-rating creates an incentive to select favorable risks. Since its current risk adjustment formula fails to neutralize the incentive for risk selection, Switzerland will complement it in 2012 with the adjuster, "Hospitalization of more than three days or living in a nursing home during the previous year". This seemingly minor fine-tuning of regulation is shown to have a potentially fatal effect on a particular health insurer A whose payments into the risk adjustment scheme would have increased substantially between 2005 and 2007 if the new risk adjustment formula had been in effect. The reason is a low rate of hospitalization thanks to a commitment to Managed Care. Therefore, A's most likely risk management response would have been to increase recognized hospitalizations by increasing length of stay from three to four days, triggering extra payments from the risk adjustment scheme at a limited once-and-for-all cost of an extra hospital day. The cost of this change of risk management strategy would have been borne by taxpayers, employers, and patients. There is an alternative that avoids the regulatory spiral described here. Health insurers could be simply permitted to charge premiums according to estimated risk. With sufficient pressure of competition, this would boil down to "price equal to expected marginal cost" since expected future health care expenditure importantly reflects the insurer's cost of enrolling an additional customer.

Wealthy individuals can pay a high risk-based premium out of their own means. The same is true of low-income individuals who are favorable risks. The problematic group are low-income individuals who are unfavorable risks. They can be entitled to an earmarked subsidy that kicks in as soon as their premium exceeds a certain percentage of their income (see Zweifel and Breuer, 2006). In fact, the new law on health insurance of 2004 introduced such a targeted subsidy in Switzerland - without however lifting the premium regulation introduced in 1911. The consequence is an avoidable fine tuning of health insurance regulation with its unhealthy impacts on not only an individual insurer but also the economy as a whole.

Bibliography

Allgood, G., Bachman, M., 2006. Effects of specialization on treatment and outcomes in screen-detected breast cancers in Wales: Cohort study. British Journal of Cancer 94, 36–42.

Amelung, V. E., 2007. Managed Care: Neue Wege im Gesundheitsmanagement - (New options in health care management).

Andreev, E., 1982. The measure of components in the analysis of length of life. Vestnik Statistiki 9, 42–47.

Arrow, K. J., 1970. Essays in the theory of risk bearing. North-Holland, Amsterdam.

Auster, R., Leveson, I., Saracheck, D., 1969. The production of health, an exploratory study. The Journal Of Human Resources 4 (4), 411–436.

Baicker, K., Chandra, A., 2004. The productivity of physician specialization: Evidence from the medicare program. American Economic Review, Papers and Proceedings W4, 184–197.

Baicker, K., Chandra, A., 2010. Understanding agglomerations in health care. In: Glaeser, E. (Ed.), The Economics of Agglomeration. University of Chicago Press, pp. 287–299.

Baranzini, A., Luzzi, G., 2001. The economic value of risks to life: evidence from the Swiss labor market. Swiss Journal of Economics and Statistics 137 (2), 149–170.

Baur, R., Andreas, H., Wieseler, S., 2001. Health care systems and reform efforts, an international comparison. Jan Böcken, Martin Butzlaff and Andreas Esche (eds.), Reforming the Health Sector, Gütersloh, 23–136.

Beck, K., 2004. Risiko Krankenversicherung, Risikomanagement in einem regulierten Krankenversicherungsmarkt (Risk Health Insurance. Risk Management in a Regulated Market for Health Insurance). Haupt Verlag, Bern.

Beck, K., Trottmann, M., Kaeser, U., Keller, B., von Rotz, S., Zweifel, P., 2006. Nachhaltige Gestaltung des Risikoausgleichs in der Schweizer Krankenversicherung - (Sustainable Design of Risk Adjustment in Swiss Health Insurance). Ott Verlag, Basel.

Becker, G. S., Murphy, K. M., 1992. The division of labor, coordination costs, and knowledge. The Quarterly Journal of Economics 107 (4), 1137–60.

Becker, K., Braendle, A., Zweifel, P., 2007. Das discrete-choice-experiment (The discrete-choice-experiment). In: Neue Versorgungsmodelle im Gesundheitswesen, Gütersloh. Jan Böcken, Katharina Janus, Uwe Schwenk and Peter Zweifel (eds.).

Bertrand, M., Duflo, E., Mullainathan, S., Mar. 2002. How much should we trust differences-in-differences estimates? NBER Working Papers 8841, National Bureau of Economic Research, Inc.

Bindman, A. B., Chattopadhyay, A., Osmond, D. H., Huen, W., Bacchetti, P., 2005. The impact of medicaid managed care on hospitalizations for ambulatory care sensitive conditions. Health Services Research 40 (1), 19 – 37.

Birkmeyer, J., Siewers, A., Finlayson, E., Stukel, T., Lucas, L., Batista, I., Welch, G., Wennberg, 2002. Hospital volume and surgical mortality in the United States. The New England Journal of Medicine 346 (15), 1128–1137.

Birkmeyer, J., Stukel, T., Siewers, A., Goodney, P., Wennberg, D., Lucas, F., 2003. Surgeon volume and operative mortality in the United States. The New England Journal of Medicine 349 (22), 2117–2127.

Browne, M. J., Hoyt, R. E., 1995. Economic and market predictors of insolvencies in the property-liability industry. The Journal of Risk and Insurance 62 (2), 309–327.

Busse, R., 2004. Disease management programs in Germany's statutory health insurance system - a gordian solution to the adverse selection of chronically ill people in competitive markets? Health Affairs 23 (3), 56–67.

Calfo, S., 2009. Medicare risk adjustment. Center for Medicare and Medicaid Services. URL `www.maac-actuary.org/`

Cameron, T., DeShazo, J., Stiffler, P., 2010. Demand for health risk reductions: A cross-national comparison between the u.s. and canada. Journal of Risk and Uncertainty 41, 245–273.

Chandra, A., Staiger, D., 2007. Productivity spillovers in health care: Evidence from the treatment of heart attacks. Journal of Political Economy 115 (1), 103–139.

Coleman, J., Katz, E., Menzel, H., 1957. The diffusion of an innovation among physicians. Sociometry 20 (4), 253–270.

Coleman, M., Quaresma, M., Berrino, F., Lutz, J., De Angelis, R., Capocacia, 2008. Cancer survival in five continents: A worldwide population-based study (concord). Lancet Oncology 9 (8), 730–756.

Cutler, D., McClellan, M., 2001. Is technological change in medicine worth it. Health Affairs 20 (5), 11–29.

Cutler, D., McClellan, M., Newhouse, J. P., 2000. How does managed care do it? Rand Journal of Economics 31 (3), 526–548.

den Exter, A., Hermans, H., Dosljak, M., Busse, R., 2004. Health Care Systems in Transition: Netherlands, Copenhagen, WHO Regional Office for Europe on behalf of the European Observatory on Health systems and Policy.

Douven, R., M. E., Pomp, M., 2007. Health care reform in the Netherlands. Die Volkswirtschaft 3, 31–33.

Douven, R., 2007. Morbidity-based risk adjustment in the Netherlands. In: Wille, V., Schneider, U. (Eds.), Wettbewerb und Risikostrukturausgleich im internationalen Vergleich; Erfahrungen aus den USA, der Schweiz, den Niederlanden und Deutschland - (Competition and risk adjustment; experiences from the US, Switzerland, the Netherlands, and Germany). Nomos, pp. 161–202.

Durbin, J., 1954. Errors in variables. Review of the International Statistical Institute 22, 23–32.

Edwards, R. D., 2008. The cost of uncertain life span. NBER Working Papers (14093). URL `http://ideas.repec.org/p/nbr/nberwo/14093.html`

Ellis, R., Van de Ven, W., 2000. Risk adjustment in competitive health plan markets. In: Culyer, A., Newhouse, J. (Eds.), Handbook of Health Economics. North-Holland, Amsterdam, pp. 755–845.

Enthoven, A., 1980. Health Plan: The Only Practical Solution to the Soaring Cost of Medical Care. Addison-Wesley, Reading, MA.

EPA, 2000. A report from the environmental economics advisory committee on EPA's white paper 'valuing the benefits of fatal cancer risk reduction. Washington D.C.

Eugster, P., Sennhauser, M., Zweifel, P., 2010. Capping risk adjustment? Journal of Health Economics 29 (4), 499–507.

Federal Council of Switzerland, 2003. Verordnung über die Krankenversicherung (Regulation on health insurance).
URL `http://www.admin.ch/ch/d/sr/832102/`

Federal Council of Switzerland, 2007. Statistik der obligatorischen Krankenversicherung 2006 - (Statistics of the statutory health insurance 2007).

Feinstein, A., Sosin, D., Wells, C., 1985. The will rogers phenomenon: Stage migration and new diagnostic techniques as a source of misleading statistics for cancer survival. New England Journal of Medicine 312, 1604–1608.

Felder, S., 2003. Managed care: Low reputation but most effective. DICE-Report Journal for Institutional Comparisons 1 (3), 15–19.

Finsterwald, D., 2004. Managed Care - Pionierland Schweiz (Managed Care - Pioneer country Switzerland). Vol. 75.

Friedman, B., 1974. Risk aversion and the consumer choice of health insurance option. The Review of Economics and Statistics 56 (2), pp. 209–214.

Fries, J., 1980. Aging, natural death, and the compression of morbidity. New England Journal of Medicine 303, 130–135.

Fuchs, V., 2004. More variation in use of care, more flat-of-the-curve medicine. Health Affairs (October 7), 104–107.

Gerdtham, U.-G., Jönsson, B., 1991. Conversion factor instability in international comparisons of health care expenditure. Journal of Health Economics 10 (2), 227 – 234.

Gillis, C., Hole, D., 1996. Survival outcome of care by specialist surgeons in breast cancer: A study of 3786 patients in the west of Scotland. BMJ 312, 145–148.

Glazer, J., McGuire, T. G., 2002. Setting health plan premiums to ensure efficient quality in health care: Minimum variance optimal risk adjustment. Journal of Public Economics 84, 153–173.

Greiner, W., 2005. Health economic evaluation of disease management programs - the german example. European Journal of Health Economics 6 (3), 191–196.

Halm, E., Lee, C., Chassin, M., 2002. Is volume related to outcome in health care? a systematic review and methodologic critique of the literature. Annals of Internal Medicine 137, 511 – 520.

Hanada, K., 1983. A formula of gini's concentration ratio and its application to life tables. Journal of Japan Statistical Society 13 (2), 95–98.

Hausman, J., 1978. Specification tests in econometrics. Econometrica 46 (6), 1251–1271.

Hoelscher, A., 2001. Oesophaguskarzinom - Operative Theorie in Zentren (Esophageal cancer - operative therapy through high volume centers). Deutsches Aerzteblatt 98 (28-29), A1890–A1894.
URL http://www.aerzteblatt.de/v4/archiv/pdf.asp?id=28041

Holly, A., Gardiol, L., Eggli, Y., Yalcin, T., 2003. Health-based risk adjustment in Switzerland: An exploration using medical information from prior hospitalization. Mimeo; Institut d'économie et management de la santé: Lausanne.

Horner, M., Ries, L., Krapcho, M., Neyman, N., Aminou, R., N., H., Altekruse, S., Feuer, E., Huang, L., Mariotto, A., Miller, B., Lewis, D., Eisner, M., Stinchcomb, D., Edwards, B., 2008. SEER Cancer Statistics Review, 1975-2006. National Cancer Institute.
URL http://seer.cancer.gov/csr/1975_2006/

Human Mortality Database, 2008. University of California Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). available at www.mortality.org.

Jack, W., 2006. Optimal risk adjustment with adverse selection and spatial competition. Journal of Health Economics 25, 908–926.

Jacobs, K., Schulze, S., 2006. Vertragswettbewerb: Mehr Freiheit bei der Partnerwahl (Competition in health insurance: More freedom of choice). Gesundheit und Gesellschaft 7, 36–41.

Joint Organization KVG, 2007. Geschäftsbericht (Business report). Solothurn.
URL www.kvg.org/info/gbericht/default.htm

Kannisto, V., 2000. Measuring the compression of mortality. Demographic Research 3 (6).

Kominski, G. F., 2007. Medicare's use of risk adjustment. National Health Policy Forum.

Krugman, P., 1991. Geography and Trade. MIT Press, Cambridge.

Laffont, J.-J., 1999. the Economics of uncertainty and information, 6th Edition. The MIT press, Cambridge MA.

Lamber, P., Aronson, J., 1993. Inequality decomposition analysis and the Gini coefficient revisited. Economic Journal 103 (420), 1221–1227.

Lamers, L., 1999. Risk adjusted capitation based on the diagnostic cost group model: An empirical evaluation with health survey information. Health Services Research 33 (6), 1727–44.

Lamers, L., Van Vliet, R., 2003a. Health-based risk adjustment: Improving the pharmacy-based cost group model to reduce gaming possibilities. European Journal of Health Economics 4, 107–114.

Lamers, L., Van Vliet, R., 2003b. The Pharmacy-based Cost Group Model: Validating and adjusting the classification of medications for chronic conditions to the Dutch situation. Health Policy 68, 113–128.

Lehmann, H., Zweifel, P., 2004. Innovation and risk selection in deregulated social health insurance. Journal of Health Economics 23, 997–1012.

Levhari, D., Mirman, L. J., 1977. Savings and consumption with an uncertain horizon. Journal of Political Economy 85 (2), 265–281.

Lichtenberg, F., 2007a. The impact of new drugs on US longevity and medical expenditure, 1990-2003: Evidence from longitudinal, disease-level data. American Economic Review 97 (2), 438–443.

Lichtenberg, F., 2007b. Pharmaceutical innovation and US cancer survival, 1992-2003: Evidence from linked SEER-MEDSTAT data. Forum for Health Economics & Policy 10 (1).
URL `http://www.bepress.com/fhep/10/1/1`

Lichtenberg, F., 2007c. Why has longevity increased more in some states than in others? Medical Progress Reports.
URL `http://www.manhattan-institute.org/html/mpr_04.htm`

Lorenz, M., 1970. Methods of measuring the concentration of wealth. Journal of American Statistical Association 9, 209–219.

Manning, W., Leibowitz, A., Goldberg, G., Rogers, W., Newhouse, J., 1984. A controlled trial of the effect of a prepaid group practice on the use of services. New England Journal of Medicine 310 (23), 1505–1510.

Miller, R., Frech, H., 2000. Is there a link between pharmaceutical consumption and improved health in OECD countries. Pharmacoeconomics 18, 33–45.

Miller, R., Luft, H., 1997. Does Managed Care lead to better or worse quality of care? Health Affairs 16 (5), 7–25.

Murphy, K. M., Topel, R. H., 2006. The value of health and longevity. The Journal of Political Economy 114 (5).

Palumbo, M., 1999. Uncertain medical expenses and precautionary saving near the end of the life cycle. Review of Economic Studies 66, 395–421.

Peltzman, S., 2009. Mortality inequality. Journal of Economic Perspectives 23 (4), 175–90.

Pope, G. C., Ellis, R. P., Ash, A. S., Liu, C.-F., Ayanian, J. Z., Bates, D. W., Burstin, H., Iezzoni, L. I., Ingber, M., 2000. Principal inpatient diagnostic cost group model for medicare risk adjustment. Health Care Financing Review 21 (3), 93–118.

Pope, G. C., Kautter, J., Ellis, R. P., Ash, A. S., Ayanian, J. Z., Iezzoni, L. I., Ingber, M., Levy, J., Robst, J., 2004. Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financing Review 25 (4), 119–141.

Pulido, J., Vierkant, R., Olson, J., Abrey, L., Schiff, D., O'Neill, P., 2009. Racial differences in primary central nervous system lymphoma incidence and survival rates. Neuro Oncology 11 (3), 318–322.

santésuisse, 2010a. Von den Erwachsenen gewählte Franchisestufen 1999-2008 - (Deductibles of the adult insured 1999-2008). Website, santésuisse Datapool. URL `http://www.santesuisse.ch/de/dyn_output_graphic_detail.html?content.`

santésuisse, 2010b. Von den Versicherten gewählte Modelle mit eingeschränkter Arztwahl 1999-2008 - (Managed care contracts 1999-2008). Website, santésuisse

Datapool.
URL http://www.santesuisse.ch/de/dyn_output_graphic_detail.html?content.

Schreyoegg, J., Stargardt, T., Velasco-Garrido, M., Busse, R., 2005. Defining the health benefit basket in nine european countries. European Journal of Health Economics 6, 2–10.

Schut, F., de Ven, W. V., 2005. Rationing and competition in the dutch health care system. Health Economics 14, 59–74.

Sfekas, A., 2009. Learning, forgetting, and hospital quality: an empirical analysis of cardiac procedures in maryland and arizona. Health Economics 18 (6), 697–711.

Shaw, J., Horrace, W., Vogel, R., 2005. The determinants of life expectancy: An analysis of the OECD health data. Southern Economic Journal 71 (4), 768–783.

Shkolnikov, V., Andreev, E., Begun, A., 2003. Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples. Demographic Research 8, 305–357.

Spycher, S., 2000. Reform des Risikoausgleichs in der Krankenversicherung (Reform of risk adjustment in health insurance). In: Bundesamt für Sozialversicherung (Ed.), Beiträge zur sozialen Sicherheit. Berne.

Stata, 2007. Stata Manual 10. Stata Press publication, College Station, Texas.

SwissDRG, 2009. Swissdrg. Www.swissdrg.org.

Le Grand, J., 1987. Inequalities in health: Some international comparisons. European Economic Review 31, 182–191.

Santésuisse, 2009. Die Reserven in der obligatorischen Krankenpflegeversicherung (Reserves in statutory health insurance). Swiss Association of Health Insurers.
URL http://www.santesuisse.ch/de/

The Federal German Ministry of Health, 2007. Die Gesundheitsreform: Infographiken - (The health care reform: Information graphs).

The Federal German Ministry of Justice, 2006. Soziales fünftes Gesetzbuch - Gesetzliche Krankenversicherung - (The statutory health insurance).

Thornton, J., 2002. Estimating a health production function for the US: Some new evidence. Applied Economics 34, 59–62.

Van de Ven, W., Schut, F., 2007. Risk equalization in an individual health insurance market: The only escape from the tradeoff between affordability, efficiency and selection - the Netherlands as a case study. Working Paper.

Van de Ven, W., Van Vliet, R., Lamers, L., 2004. Health-adjusted premium subsidies in the Netherlands. Health Affairs 23 (3), 45–55.

Van de Ven, W. P., Schut, F. T., 2008. Universal mandatory health insurance in the Netherlands: A model for the United States? Health Affairs 27 (3), 771–781.

Wiechmann, M., 2003. Managed Care: Grundlagen, internationale Erfahrungen und Umsetzung im deutschen Gesundheitswesen (Managed Care: Basic concepts, international expericence and implementation in the German health care system).

Wilmoth, J., Horiuchi, S., 1999. Rectangularization revisited: Variability of age at death within human population. Demography 36 (4), 475–495.

Zelen, M., 1976. Theory of early detection of breast cancer in the general population. In: Reusen, J.-C., Rozencweig, M. (Eds.), Trends in Research and Treatment. Raven Press, New York, pp. 287–299.

Zweifel, P., 1998. Managed Care in Germany and Switzerland: Two approaches to a common problem. Pharmacoeconomics 1, 1–8.

Zweifel, P., 2006. Building a competitive insurance system: Switzerland's strategy for managed care-healthcare. Pharmacoeconomics 24, 109–117.

Zweifel, P., Breuer, M., 2006. The case for risk-based premiums in public health insurance. Health Economics, Policy and Law 1 (2), 171–188.

Zweifel, P., Breyer, F., Kifmann, M., 2009. Health Economics, 2nd Edition. Springer Verlag, Boston MA.

Zweifel, P., Eisen, R., 2003. Versicherungsökonomie (Insurance Economics). Springer Verlag, 2nd. ed., Heidelberg.

Zweifel, P., Steinmann, L., Eugster, P., 2005. The Sisyphus syndrome in health revisited. International Journal of Health Care Finance and Economics 5, 127–145.

Zweifel, P., Telser, H., Vaterlaus, S., 2006. Consumer resistance against regulation: The case of health care. Journal of Regulatory Economics 29, 319–332.